```
   1 TGGCAGTGGG CGGCGTAGAG CACTGCAGCA GCAATGACGG AGGGCACGTG
  51 TCTGCGGCGC CGAGGGGGCC CCTACAAGAC CGAGCCCGCC ACCGACCTCG
 101 GCCGCTGGCG ACTCAACTGC GAGAGGGGCC GGCAGACGTG GACCTACCTG
 151 CAGGACGAGC GCGCCGGCCG CGAGCAGACC GGCCTGGAAG CCTACGCCCT
 201 GGGGCTGGAC ACCAAGAATT ACTTTAAGGA CTTGCCCAAA GCCCACACCG
 251 CCTTTGAGGG GGCTCTGAAC GGGATGACAT TTTACGTGGG GCTGCAGGCT
 301 GAGGATGGGC ACTGGACGGG TGATTATGGT GGCCCACTTT TCCTCCTGCC
 351 AGGCCTCCTG ATCACTTGCC ACGTGGCACG CATCCCTCTG CCAGCCGGAT
 401 ACAGAGAAGA GATTGTGCGG TACCTGCGGC ACATTGAGGA TAAGTCCACC
 451 GTGTTTGGGA CTGCGCTCAA CTATGTGTCT CTCAGAATTC TGGGTGTTGG
 501 GCCTGACGAT CCTGACCTGG TACGAGCCCG GAACATTCTT CACAAGAAAG
 551 GTGGTGCTGT GGCCATCCCC TCCTGGGGGA AGTTCTGGCT GGCTGTCCTG
 601 AATGTTTACA GCTGGGAAGG CCTCAATACC CTGTTCCCAG AGATGTGGCT
 651 GTTTCCTGAC TGGGCACCGG CACACCCCTC CACACTCTGG TGCCACTGCC
 701 GGCAGGTGTA CCTGCCCATG AGCTACTGCT ACGCCGTTCG GCTGAGTGCC
 751 GCGGAAGACC CGCTGGTCCA GAGCCTCCGC CAGGAGCTCT ATGTGGAGGA
 801 CTTCGCCAGC ATTGACTGGC TGGCGCAGAG GAACAACGTG GCCCCCGACG
 851 AGCTGTACAC GCCGCACAGC TGGCTGCTCC GCGTGGTATA TGCGCTCCTC
 901 AACCTGTATG AGCACCACCA CAGTGCCCAC CTGCGGCAGC GGGCCGTGCA
 951 GAAGCTGTAT AACACATTG TGGCCGACGA CCGATTCACC AAGAGCATCA
1001 GCATCGGCCC GATCTCGAAA ACCATCAACA TGCTTGTGCG CTGGTATGTG
1051 GACGGGCCCG CCTCCACTGC CTTCCAGGAG CATGTCTCCA GAATCCCGGA
1101 CTATCTCTGG ATGGGCCTTG ACGGCATGAA AATGCAGGGC ACCAACGGCT
1151 CACAGATCTG GGACACCGCA TTCGCCATCC AGGCTCTGCT TGAGGCGGGC
1201 GGGCACCACA GGCCCGAGTT TTCGTCCTGC CTGCAGAAGG CTCATGAGTT
1251 CCTGAGGCTC TCACAGGTCC CAGATAACCC TCCCGACTAC CAGAAGTACT
1301 ACCGCCAGAT GCGCAAGGGT GGCTTCTCCT TCAGTACGCT GGACTGCGGC
1351 TGGATCGTTT CTGACTGCAC GGCTGAGGCC TTGAAGGCTG TGCTGCTCCT
1401 GCAGGAGAAG TGTCCCCATG TCACCGAGCA CATCCCCAGA GAACGGCTCT
1451 GCGATGCTGT GGCTGTGCTG CTGAACATGA GAAATCCAGA TGGAGGGTTC
1501 GCCACCTATG AGACCAAGCG TGGGGGGCAC TTGCTGGAGC TGCTGAACCC
1551 CTCGGAGGTC TTCGGGGACA TCATGATTGA CTACACCTAT GTGGAGTGCA
1601 CCTCAGCCGT GATGCAGGCG CTTAAGTATT TCCACAAGCG TTTCCCGGAG
1651 CACAGGGCAG CGGAGATCCG GGAGACCCTC ACGCAGGGCT TAGAGTTCTG
1701 TCGGCGGCAG CAGAGGGCCG ATGGCTCCTG GGAAGGCTCC TGGGGAGTTT
1751 GCTTCACCTA CGGCACCTGG TTTGGCCTGG AGGCCTTCGC CTGTATGGGG
1801 CAGACCTACC GAGATGGGAC TGCCTGTGCA GAGGTCTCCC GGGCCTGTGA
1851 CTTCCTGCTG TCCCGGCAGA TGGCAGACGG AGGCTGGGGG GAGGACTTTG
1901 AGTCCTGCGA GGAGCGGCGT TATGTGCAGA GTGCCCAGTC CCAGATCCAC
1951 AACACATGCT GGGCCATGAT GGGGCTGATG GCCGTTCGGC ATCCTGACAT
2001 CGAGGCCCAG GAGAGAGGAG TCCGGTGTCT ACTTGAGAAA CAGCTCCCCA
2051 ATGGCGACTG GCCGCAGGAA AACATTGCTG GGGTCTTCAA CAAGTCCTGT
2101 GCCATCTCCT ACACGAGCTA CAGGAACATC TTCCCCATCT GGGCCCTCGG
2151 CCGCTTCTCC CAGCTGTACC CTGAGAGAGC CCTTGCTGGC CACCCCTGAG
2201 AACATGCCTA CCTGCTGGGT GCCGTCTGTG CGTTCCAGTG AGGCCAAGGG
2251 GTCCTGGCCG GGTTGGGGAG CCCTCCCATA ACCCTGTCTT GGGCTCCAAC
2301 CCCTCAACCT CTATCTCATA GATGTGAATC TGGGGGCCAG GCTGGAGGCA
2351 GGGATGGGGA CAGGGTGGGT GGCTTAGACT CTTGATTTTT ACTGTAGGTT
2401 CATTTCTGAA AGTAGCTTGT CGGGCTTGGG TGAGGAAGGG GGCACAGGAG
2451 CCGTGACCCC TGAGGAGGCA CAGCGCCTTC TGCCACCTCT GGGCACGGCC
2501 TCAAGGTAGT GAGGCTAGGA GGTTTTTTCT GACCAATAGC TGAGTTCTTG
2551 GGAGAGGAGC AGCTGTGCCT GTGTGATTCC TTAGTGTCGA GTGGGCTCTG
2601 GGCTGGGGTC GGCCCTGGGC AGGCTTCTCC TGCACCTTTT GTCTGCTGGG
2651 CTGAGGGACA CGAGGGCAAC CCTGTGACAA TGGCAGGTAG TGTGCATCCG
2701 TGAATAGCCC AGTGCGGGGG TTGCTCATGG AGCATCCTGA GGCCGTGCAG
2751 CAGGGAGCCC CATGCCCCTG GGTCGTGAGC TTGCCTGCGT ATGGGGTGGT
2801 GTCATGGAGC CTCATGCCCC TGGGTCGTGA GCTCGCCTGA GTATGGGGTG
2851 GTGTCATGGA GCCGCATACC CCTGGGTTGT GAGCTCGCCT GCATATGCAG
2901 GGTCTGTCAT GGAACATCCC AAGTCTGTGC AGCAGGGAGC CCCATGCCCC
2951 TGGGACATGA ACCCACCTGC GTGGAATGCT GTTTGTGAGG TGTCTACAGG
3001 GTTTATAGTA GTCTTGTGGA CACAGAAATG CACAGGGGAC ACTTACGGAC
3051 ACAGAAATGC ACAGGGGAGG CCGAGCATAA CCAGGGGTGA GGGGCAGGCA
3101 GCAGTTGTAG TTACTGCCGC GGGGCACTGC TATGTGCAGG GACAGCCAGC
```

FIGURE 1A

```
3151 GCCCAGCCCA TCACCACTCC CTGGGCTGGC TGGCAGGTAT GGCACCCTGG
3201 GAGCCCGGCA TATACCCAGG GCACCCCTAC GGCTGCCGCC AGTCTCATGC
3251 CCAGGTGGGT GCTCTGGGCT GGAGCGAGGG CCAGGTTTTG GGCCGAGGCT
3301 TCCCCAGGCA ATCCTGTGAG CTCCCTTCTA GCCTCTGACC CAGTCTGGTC
3351 TGGCTTGCAT GGATGTAGGG CTTGGGGTGG GAAGTTCAGG TCCTGGCTTT
3401 GCCTTTGCCT GATGTGGATG AGCAGCTCAC ATGCTCAGGG CCACCTGAGA
3451 CTGTCACTGC TCTCCCCTGG CTACTGGGAG GAGTCACTGA GAGCTTCGTT
3501 ACCCCTGCTG CCTTGCCCAG GGCACACCCT ATACCTCCTC ATCTGCTCTT
3551 CCCCTCCCTG CCGCCTTCTG GGCAGGTAGC AGTCCCTGGC CTCTCCCCCT
3601 GGCTGATCAC TCTCCCTCAG GCAGTGGAGA TCTGCGTCTG GACACCCTCA
3651 GATCCTGTCA TTGCCTGCCC AGAGTCCTTC AGGGGCACCC CTCTGCCTTG
3701 GTGTGCGGTC CAGGGCTCTC ACCCAGGTGC CGCACCCTCT GGGGTCTTCT
3751 GTCCAGCTCC CTTGCCCCAT GTGCTGTCAC TGACTCTCCT TGGGACTCGC
3801 CTGCCTGCTC AGAGCCCTGC AGGGCTTGGT CAGCTGCCTG TTCAGTGTCA
3851 ACACTTCCCT GCACATCTTA AAACTGGGCT TTATTTTCGC TGAAGGAACT
3901 GTGTTGGGAC CCTTGACATC TGTCAGGTTT GCACATGCTG TTTTTTTTTC
3951 TCAGCCCACG TGTTCTCCCC CACGTGGGGT AGCAGCAGGA CAGACAGTGA
4001 ATCACAGAGT CTGCCCTGAG CAGAGGCTGC TGTCCCTGGG ACTCCTAGCC
4051 ATGGTCAGAC TGTACAAAAC GGTTTTCCAG AAATGAAATG TAAATCCATT
4101 TTTATACTGA AAATGTTACT GAAAGTCACT TTTATGAGCA TCTGCCTTAA
4151 TAAACAGACA TTGATTCCCT TAAAAAAAAA AAAAAAAAAA AAAAAAAAAA
4201 AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAA   (SEQ ID NO:1)
```

FEATURES:
5'UTR: 1-33
Start Codon: 34
Stop Codon: 2197
3'UTR: 2200

Homologous proteins:
Top 10 BLAST Hits

	Score	E
CRA|18000005000949 /altid=gi|4505027 /def=ref|NP_002331.1| lano...	1530	0.0
CRA|18000005227733 /altid=gi|4808278 /def=emb|CAB42828.1| (AJ23...	1524	0.0
CRA|18000005013642 /altid=gi|1098635 /def=gb|AAA91023.1| (U3135...	1315	0.0
CRA|18000004977416 /altid=gi|1352388 /def=sp|P48450|ERG7_RAT LA...	1305	0.0
CRA|18000005002424 /altid=gi|984145 /def=emb|CAA61078.1| (X8780...	1224	0.0
CRA|100000004433519 /altid=gi|8886139 /def=gb|AAF80384.1|AF1599...	689	0.0
CRA|335001098658178 /altid=gi|11279144 /def=pir||T48782 lanoste...	611	e-173
CRA|18000005223063 /altid=gi|4589852 /def=dbj|BAA76902.1| (AB02...	609	e-173
CRA|18000005171896 /altid=gi|3688598 /def=dbj|BAA33460.1| (AB00...	607	e-172
CRA|1000682333668 /altid=gi|6045133 /def=dbj|BAA85266.1| (AB033...	605	e-172

BLAST dbEST hits:

	Score	E
gi|10993792 /dataset=dbest /taxon=96...	1538	0.0
gi|10159427 /dataset=dbest /taxon=96...	1358	0.0
gi|9340844 /dataset=dbest /taxon=960...	1108	0.0
gi|11251687 /dataset=dbest /taxon=96...	1065	0.0
gi|11258382 /dataset=dbest /taxon=96...	995	0.0
gi|10322370 /dataset=dbest /taxon=96...	910	0.0

FIGURE 1B

EXPRESSION INFORMATION FOR MODULATORY USE:

library source:

From BLAST dbEST hits:

gi|10993792 teratocarcinoma
gi|10159427 ovary
gi|9340844 uterus
gi|11251687 muscle
gi|11258382 brain
gi|10322370 colon

From tissue screening panels:

hippocampus

FIGURE 1C

```
  1 MTEGTCLRRR GGPYKTEPAT DLGRWRLNCE RGRQTWTYLQ DERAGREQTG
 51 LEAYALGLDT KNYFKDLPKA HTAFEGALNG MTFYVGLQAE DGHWTGDYGG
101 PLFLLPGLLI TCHVARIPLP AGYREEIVRY LRHIEDKSTV FGTALNYVSL
151 RILGVGPDDP DLVRARNILH KKGGAVAIPS WGKFWLAVLN VYSWEGLNTL
201 FPEMWLFPDW APAHPSTLWC HCRQVYLPMS YCYAVRLSAA EDPLVQSLRQ
251 ELYVEDFASI DWLAQRNNVA PDELYTPHSW LLRVVYALLN LYEHHHSAHL
301 RQRAVQKLYE HIVADDRFTK SISIGPISKT INMLVRWYVD GPASTAFQEH
351 VSRIPDYLWM GLDGMKMQGT NGSQIWDTAF AIQALLEAGG HHRPEFSSCL
401 QKAHEFLRLS QVPDNPPDYQ KYYRQMRKGG FSFSTLDCGW IVSDCTAEAL
451 KAVLLLQEKC PHVTEHIPRE RLCDAVAVLL NMRNPDGGFA TYETKRGGHL
501 LELLNPSEVF GDIMIDYTYV ECTSAVMQAL KYFHKRFPEH RAAEIRETLT
551 QGLEFCRRQQ RADGSWEGSW GVCFTYGTWF GLEAFACMGQ TYRDGTACAE
601 VSRACDFLLS RQMADGGWGE DFESCEERRY VQSAQSQIHN TCWAMMGLMA
651 VRHPDIEAQE RGVRCLLEKQ LPNGDWPQEN IAGVFNKSCA ISYTSYRNIF
701 PIWALGRFSQ LYPERALAGH P  (SEQ ID NO:2)
```

FEATURES:
Functional domains and key regions:
[1] PDOC00001 PS00001 ASN_GLYCOSYLATION
N-glycosylation site

Number of matches: 2
```
        1    371-374  NGSQ
        2    686-689  NKSC
```

[2] PDOC00005 PS00005 PKC_PHOSPHO_SITE
Protein kinase C phosphorylation site

Number of matches: 5
```
        1    149-151  SLR
        2    247-249  SLR
        3    149-151  SLR
        4    247-249  SLR
        5    494-496  TKR
```

[3] PDOC00006 PS00006 CK2_PHOSPHO_SITE
Casein kinase II phosphorylation site

Number of matches: 7
```
        1     49-52   TGLE
        2     72-75   TAFE
        3    238-241  SAAE
        4    434-437  STLD
        5    518-521  TYVE
        6    591-594  TYRD
        7    624-627  SCEE
```

FIGURE 3A

[4] PDOC00008 PS00008 MYRISTYL
N-myristoylation site

Number of matches: 11
```
        1       76-81 GALNGM
        2      107-112 GLLITC
        3      142-147 GTALNY
        4      173-178 GGAVAI
        5      369-374 GTNGSQ
        6      487-492 GGFATY
        7      552-557 GLEFCR
        8      564-569 GSWEGS
        9      571-576 GVCFTY
       10      577-582 GTWFGL
       11      595-600 GTACAE
```

[5] PDOC00825 PS01074 TERPENE_SYNTHASES
Terpene synthases signature

```
          563-577 DGSWEGSWGVCFTYG
```

Membrane spanning structure and domains:

Helix	Begin	End	Score	Certainty
1	95	115	1.321	Certain
2	173	193	0.944	Putative
3	569	589	1.311	Certain

FIGURE 3B

```
>CRA|18000005000949 /altid=gi|4505027 /def=ref|NP_002331.1|
          lanosterol synthase (2,3-oxidosqualene-lanosterol
          cyclase); Lanosterol synthase; human lanosterol synthase
          [Homo sapiens] /org=Homo sapiens /taxon=9606
          /dataset=nraa /length=732
        Length = 732
```

```
 Score = 1530 bits (3917), Expect = 0.0
 Identities = 720/732 (98%), Positives = 721/732 (98%), Gaps = 11/732 (1%)
```

```
Query:   1  MTEGTCLRRRGGPYKTEPATDLGRWRLNCERGRQTWTYLQDERAGREQTGLEAYALGLDT 60
            MTEGTCLRRRGGPYKTEPATDLGRWRLNCERGRQTWTYLQDERAGREQTGLEAYALGLDT
Sbjct:   1  MTEGTCLRRRGGPYKTEPATDLGRWRLNCERGRQTWTYLQDERAGREQTGLEAYALGLDT 60

Query:  61  KNYFKDLPKAHTAFEGALNGMTFYVGLQAEDGHWTGDYGGPLFLLPGLLITCHVARIPLP 120
            KNYFKDLPKAHTAFEGALNGMTFYVGLQAEDGHWTGDYGGPLFLLPGLLITCHVARIPLP
Sbjct:  61  KNYFKDLPKAHTAFEGALNGMTFYVGLQAEDGHWTGDYGGPLFLLPGLLITCHVARIPLP 120

Query: 121  AGYREEIVRYLR----------HIEDKSTVFGTALNYVSLRILGVGPDDPDLVRARNIL 169
            AGYREEIVRYLR          HIEDKSTVFGTALNYVSLRILGVGPDDPDLVRARNIL
Sbjct: 121  AGYREEIVRYLRSVQLPDGGWGLHIEDKSTVFGTALNYVSLRILGVGPDDPDLVRARNIL 180
                                   · ·:
Query: 170  HKKGGAVAIPSWGKFWLAVLNVYSWEGLNTLFPEMWLFPDWAPAHPSTLWCHCRQVYLPM 229
            HKKGGAVAIPSWGKFWLAVLNVYSWEGLNTLFPEMWLFPDWAPAHPSTLWCHCRQVYLPM
Sbjct: 181  HKKGGAVAIPSWGKFWLAVLNVYSWEGLNTLFPEMWLFPDWAPAHPSTLWCHCRQVYLPM 240

Query: 230  SYCYAVRLSAAEDPLVQSLRQELYVEDFASIDWLAQRNNVAPDELYTPHSWLLRVVYALL 289
            SYCYAVRLSAAEDPLVQSLRQELYVEDFASIDWLAQRNNVAPDELYTPHSWLLRVVYALL
Sbjct: 241  SYCYAVRLSAAEDPLVQSLRQELYVEDFASIDWLAQRNNVAPDELYTPHSWLLRVVYALL 300

Query: 290  NLYEHHHSAHLRQRAVQKLYEHIVADDRFTKSISIGPISKTINMLVRWYVDGPASTAFQE 349
            NLYEHHHSAHLRQRAVQKLYEHIVADDRFTKSISIGPISKTINMLVRWYVDGPASTAFQE
Sbjct: 301  NLYEHHHSAHLRQRAVQKLYEHIVADDRFTKSISIGPISKTINMLVRWYVDGPASTAFQE 360

Query: 350  HVSRIPDYLWMGLDGMKMQGTNGSQIWDTAFAIQALLEAGGHHRPEFSSCLQKAHEFLRL 409
            HVSRIPDYLWMGLDGMKMQGTNGSQIWDTAFAIQALLEAGGHHRPEFSSCLQKAHEFLRL
Sbjct: 361  HVSRIPDYLWMGLDGMKMQGTNGSQIWDTAFAIQALLEAGGHHRPEFSSCLQKAHEFLRL 420

Query: 410  SQVPDNPPDYQKYYRQMRKGGFSFSTLDCGWIVSDCTAEALKAVLLLQEKCPHVTEHIPR 469
            SQVPDNPPDYQKYYRQMRKGGFSFSTLDCGWIVSDCTAEALKAVLLLQEKCPHVTEHIPR
Sbjct: 421  SQVPDNPPDYQKYYRQMRKGGFSFSTLDCGWIVSDCTAEALKAVLLLQEKCPHVTEHIPR 480

Query: 470  ERLCDAVAVLLNMRNPDGGFATYETKRGGHLLELLNPSEVFGDIMIDYTYVECTSAVMQA 529
            ERLCDAVAVLLNMRNPDGGFATYETKRGGHLLELLNPSEVFGDIMIDYTYVECTSAVMQA
Sbjct: 481  ERLCDAVAVLLNMRNPDGGFATYETKRGGHLLELLNPSEVFGDIMIDYTYVECTSAVMQA 540

Query: 530  LKYFHKRFPEHRAAEIRETLTQGLEFCRRQQRADGSWEGSWGVCFTYGTWFGLEAFACMG 589
            LKYFHKRFPEHRAAEIRETLTQGLEFCRRQQRADGSWEGSWGVCFTYGTWFGLEAFACMG
Sbjct: 541  LKYFHKRFPEHRAAEIRETLTQGLEFCRRQQRADGSWEGSWGVCFTYGTWFGLEAFACMG 600

Query: 590  QTYRDGTACAEVSRACDFLLSRQMADGGWGEDFESCEERRYVQSAQSQIHNTCWAMMGLM 649
            QTYRDGTACAEVSRACDFLLSRQMADGGWGEDFESCEERRY+QSAQSQIHNTCWAMMGLM  ·
Sbjct: 601  QTYRDGTACAEVSRACDFLLSRQMADGGWGEDFESCEERRYLQSAQSQIHNTCWAMMGLM 660

Query: 650  AVRHPDIEAQERGVRCLLEKQLPNGDWPQENIAGVFNKSCAISYTSYRNIFPIWALGRFS 709
            AVRHPDIEAQERGVRCLLEKQLPNGDWPQENIAGVFNKSCAISYTSYRNIFPIWALGRFS
Sbjct: 661  AVRHPDIEAQERGVRCLLEKQLPNGDWPQENIAGVFNKSCAISYTSYRNIFPIWALGRFS 720

Query: 710  QLYPERALAGHP 721
            QLYPERALAGHP
Sbjct: 721  QLYPERALAGHP 732   (SEQ ID NO:4)
```

FIGURE 3C

Hmmer search results (Pfam):

Model	Description	Score	E-value	N
PF00432	Prenyltransferase and squalene oxidase repea	83.9	1.7e-22	3

Parsed for domains:

Model	Domain	seq-f	seq-t		hmm-f	hmm-t		score	E-value
PF00432	1/3	133	154	..	23	45	.]	6.6	3.8
PF00432	2/3	547	589	..	1	45	[]	40.1	8e-10
PF00432	3/3	599	647	..	1	45	[]	39.4	1.3e-09

FIGURE 3D

```
   1 TCATGACTGC CCCTAGAAGC TTAACTGTGT CAATTCTCAG ACGTAGTTTA
  51 CAGCTTTTTC TTTTCTTTCA GACATTAAAA AGAGCGGATT ATTTTACTCA
 101 TAAAAAGTCC AGTCCATTAA GATATCAAAA CTCAAACTCT TATCCAGTTG
 151 AAACCTCTTC CCTCACCTAG CTTTGCCAGG TTCAGTGTGA GATTCCATCC
 201 AGGCTGAAGC CCCTTATCCC TATTCTTCAT GTTTCTACAT GGAGGAACTT
 251 ACCTGGAGAA AAACTTCCAG CCTCTTTCTG CTTCCAGAGA AGTAGAGTGA
 301 CTCATTTGAT TGAATTCAG AGAACAGATA GGGTGGAGTG TGCTCAGGCT
 351 CCTCTGGGTA CTCTTTCTGG GGTCTGTGGG TTGACTGGAG GGGTGTCTTC
 401 TGGTGGGCAC TCAATTGCAT AGTGCTTGGT GAGGCAGTTT CATGGCCTAG
 451 AGGCTGGGGG ATATGTTTGT CTGACTTACG GGTGATTTAG TAGCTTGCCC
 501 TCTTGCTTGC AGATTTAAGC CTTGTCCTTC AAGCTAGGTT TTTAATTTGT
 551 GGCAAAGCTG ATATTTTGAT ACCCACCCAT CTTATTGCTG TGTCTTTTTC
 601 ATCCGTTTCT GAACTGGGAT AGGAAGAGGT GATTATCCTT GATTGTCTAA
 651 AACCCCGCTA TTCCACTGTG GGGAAGGTGC CTGTGGGTAT TCTTTTGTCC
 701 ACTCTCTCTT CCAACTTTCT CCTCCGGCTT GCTGTGGCTC ACCGCCCCTT
 751 CGAAGTTAGG CTGGGGGTAG GAATTGAGGA GTGGGTGCCG AAATGCTCAC
 801 TAGGCTGGGG CAGTTGTAAC TGGATGTCAG GGCTTCTGTG GGCCAGGTGA
 851 AGACATGCTG GGGTCTTCTG TGGGTCCTTG ACCTGACTTA GGGACCACTG
 901 GCTGCAGCCT CCAGACGTCA GCCATGTTTC CAACAGTCAG ACGCCCCCTG
 951 CCCTGTTGCG CCCGGCTGTC CCTTCCAAGT TCGGTCACTC GCTCTGCCTC
1001 CATCTTCCTC TTCCCTCTGC TGCTAAGGCT TTTCACCTTT AATTTCTCCT
1051 GGGGCCACCC CCAACTCCAG CGACCCCGTG AGCAGCTGAG GCTCTACCGC
1101 GCTCGGTCCT GGCCAGCGAC GCAGCCCTTC CCTGGCGGGG CTCCAGGGCT
1151 TCTGGCCCCT GTGGTCCGCC AGGTGTGGGG GCCCACGGCC TCACCGCGCC
1201 TACCCCACTC CCCCCGGCGA AGCTACGCGG CGCTCAGCTT CCCAGGGACG
1251 CCGGCGGCGC CCTCGGCTCC TCCGCTCCGC CCCGCCCTCC CCCTGGTCTC
1301 GCACTGGAGC CGACGGCCCG CGCCCACCTC ACCTCAGGGC GGCCTCCCGC
1351 CCCCACCCCC GGCCCGGCG TCCGGGCAAA TCCTGCAGCG CGAGAGCAAT
1401 TCCCTGCCAC CCGACCTTCG CACTCGCTGT CGCTCGCTCG AGCCTCGCTC
1451 CCCACGTCCT TCCTTCCGAC CCGCGGCTGG ACCTCCTCA CAAATTTCTC
1501 AGAGAGGCTC ACCTCAAAGC GCGGCGCACG AGGCCGGGCT CCCGGGACGC
1551 AAGCCTCTAG AGGGCGCGCG AGAGGCCCCG CCCCCGCCCT TCGGCCCCAC
1601 CCACCAGCCC CGCCCCCACC CGCACCCACC AGGCCCCGCC CCCACCTCCC
1651 CACCCACCAG CCCCGCCCCC ACCTCCCCAC CCACCAGCCC CGCCCCTCAT
1701 GCCCCGCCAA TAAGGCCCCA CCCGCCTCCC CCGTCCCGTC GCCTTCACCC
1751 ACCATCCCCG CTCCCTCAGG CCCCGCCCCA GGCACCAAGC
1801 GCTCCACCAC TGTGGTCGCC TGGCACACCC CGGGGTCACG CTCGCGGCGC
1851 TCTGATTGGT TGCGTGGGCG TCGGCCCACC TAAGCCTGAG CGCCTGCCGA
1901 GGCCTGCGCC TGCGTAGTGC GCGCGGGAGG GGCGGGAGGG GCGGGAGGGG
1951 CGGGAGGGGC GGGGCTGGGC GGCAGGTCCC GGGTGCGGAC ATCTGGCAGC
2001 TGGCAGTGGG CGGCGTAGAG CACTGCAGCA GCAATGACGG AGGGCACGTG
2051 AGTCCCCTCG CCCCGGGCTC CTGACGAATG CGGGGTGGTC CTAGGTGCTG
2101 AGGAGAGCGC GACTGGGGCA GTGGGCCGGC GGCCGGCGTT GGGGCGGGGC
2151 CTGGGTCGCT GATGGCCGGT GGTCCTCAGG TGTCTGCGGC GCCGAGGGGG
2201 CCCCTACAAG ACCGAGCCCG CCACCGACCT CGGCCGCTGG CGACTCAACT
2251 GCGAGAGGGG CCGGCAGACG TGGACCTACC TGCAGGACGA GCGCGCCGGC
2301 CGCGAGCAGA CCGGCCTGGA AGCCTACGCC CTGGGGCTGG ACACCGTAAG
2351 TTGCTTCCGC GGAGCGTCAG CGAGCTCGGG ACCCTGAGGG GTGAGCCGTG
2401 AGGAGCACGT TTTCTCTCAG AAAGGCGGGT GGGAGGACCC GGCCAGCGAC
2451 GCCCATCCCC AAGGCGAGCG CCCACGGGAA CTGCGTTCGC GGGCCCCTCC
2501 GCTTCAGCCC CTTCATCTCT AAACCACGCA TAGGAGACTC CTAATGTTTT
2551 ATTTTTTAGC ACCTTATTTT GAGATAATTT TTGACTTATA GGAGAGTTGC
2601 AAAGATAGTT GTAACTTTGT TTTTATTCAC AAAAAGTGTT TGGATCCACT
2651 GTCTTAGTTG TGTGCATTGT AAGAGATTTT GGTCGTCAGA GTCTGCAGTG
2701 TAAACAGGGT CTCCTGCCGA GCCCCGGCCA CCGAGGGAAA GGCTGTGCCG
2751 CCCCTTGGGC CCTCTTTGAG AGGCCCGAGT CCCAGGCCCA GGTCGGCACC
2801 CGTGCCCCAC CCTACAGTCT GGGTGCCTGG TTTATTCCAG ACATCTTGGA
2851 GAAGTTGTGA AGAATACATG ACTGGCAAAT AAAGCAACGA AAATGTGCAG
2901 CTGTTCTTTT ACTTTGCTGA GGTGTGATGC TCTCATCAAA GAGTTTCAGA
2951 CTTTTGATGG AAACAGCTGA AACTTTTAAA GTAATTACA TTCACTGTTT
3001 TGACTTGGGC TGTATGTGAA GAGGGTTCCT CTGGCCGGGC AACAGTCCCG
3051 TCAGCTATCT CTTTTTTTTT TTTCGATCT CTTTGCAGAA GAATTACTTT
3101 AAGGACTTGC CCAAAGCCCA CACCGCCTTT GAGGGGGCTC TGAACGGGAT
```

FIGURE 3A

```
3151 GACATTTTAC GTGGGGCTGC AGGCTGAGGA TGGGCACTGG ACGGGTGATT
3201 ATGGTGGCCC ACTTTTCCTC CTGCCAGGTA GGAGTATGCT GCCCCAGCCT
3251 GATGGTATGG CCACCCTGGA TCACCCTTGG GATCCTGGCC CAGCCTGGTC
3301 TAGGGTTTTG ATGAAGCAGG TGAAAATCCA GGGGCTCACA AGAAAAGGGC
3351 TGGCAAACTC TGCCCTATGT CAGAGTCGTC CTGCTATTGG TCTAGGGGAT
3401 CAGCTAGCCT TGCCAGTGTA GGGTGACAGG CTCTCTGATA AGAGAAGCAA
3451 GTGGTTCTCT AGGGCTCTGT GTTGCCTTGA GGGAGGAGGA AGGTGGGCTT
3501 TGAAGTCTCA GTACAGGATG GGATGGACAT TCCAGGTGGA AGGCCCAGCC
3551 TATGCCAAGG GGCTGTAGGT GGGCAGAGTG GTGGGTGGGG AGCTGATATC
3601 TGCTGTGAAC TTCCTCGGGG CTATTGCAGG AGAGCTTCAG GTTCAGGCTG
3651 GTGAGTAGGA GGAGCATAGC AGTTGGACTG CCTGGGTATT GAACTGATTT
3701 GGCTACACAA GACTATTTTG CATCCTGGGA GTGTTTCTCT ACAGAAATCC
3751 TCAGCCTTGT AAAATGGGAA ATTCCCTCCT ATGAATTTAT GCAATAGGAC
3801 TTTTTTCCCT AGTGACTTGT AATCACATTG TTTCAATGAC GTGAATTCCT
3851 ACATAAATAG GTTTTGTTTC TGTGATAACT CTTACTGATA CATCATTTTC
3901 TTTTACTACG CTGACTTTGT AATAGATAGA AAGTCCTTAT ATACCTTTGT
3951 TGCCTTTCTT TTTAAAACAT CTCTTACCTG TGTCTATTCA TTTACTCATC
4001 CAAATTGCCT TTATCCTGAT TTTGTCCCAG ACTTGAAATG AAGTTGCAAT
4051 AGGCTTATAT GTTAGTTTGG GAAGAGTTGG CCTTTAACGT TAAAAACAGT
4101 TCCATGGTGT TTACTGTAGG CCAAGCCCTG CTCAAGGCCT GTTCTTCTTT
4151 TAGTCCTTAG AATAAGCCTA ATGAGATACA TTAGAAAGCT GAGGCACATT
4201 TATTCCAGGT AACCAGACTA GCAGGAGGAG CACTGGGATC CCCATCTCTG
4251 CTTTGACTTC TAGCCCTGCT GCCACCTGGA CTGTACAGCA TTGAGTTTTT
4301 CTGTCCTGGG ATTTGAGGGC CTGTCCTTAG GGGAAGTCAA GGTGCTCTTC
4351 TTCCCTTGGC CCCATCAGGG CCTGTTTAGA CTGTTCTCAG GGCTCGTGGT
4401 AAGGCAATGA CATAGAGTTG GTCAGGAGAT GGGTCAGCCC CACTTTGCCT
4451 CTGTAGCCTG ACCTGTGACA GGATTGGAAT CAGGTTTGGT CATGTGCACA
4501 GTGTCAGGCA TGCAGTGGTG CTTGGTCAGT GGGGATTACT GTGTTGTTTG
4551 TTCTTGCTGC TTTGGCTCTG GGCTTAGCTG GCTGGGACCC TTCCTGTGGG
4601 CTGGCTGTGA GTTGGAGTTT TTTTGTATTT TTTTTTTTTT TTTGAGACAG
4651 CGTTCGCTCT TGTTCCCCAG GCTGGAGTGC AATGGCACAA TTTTGGCTCG
4701 TTGCAGCCTC TGCCTCCTGG GTGCAAGTGA TTCTCCTGCC TCAGCCTCCT
4751 GTAGGGTCCA GCCCCACAGG GTCGGTAGGT TTTTCTCCCT GTGTGCGGAG
4801 ATGAGAGATT GTAGAAATAA AGACACAAGA CAGAGAGATG AAAGAAAAGA
4851 CAGCTGGGCC CCGGGGGACC ACTACCACCA AGACGTGGAA ACCGGTAGTG
4901 GCCCTGAATG CCAGGCTGCG CTGATATTTA TTGGATACAA GACAAAGGGG
4951 CAGGGTAAGG AGTGTGAGCC ATCTCCAATG ATAGGTAAGG TCACATGGGT
5001 CACGTGTCCA CTGGACAGTG GGCCCTTCCC TGCCTGGCAG CCGAGGCAGA
5051 GAGTGGGAGA GAGAGAGAGA GAGACAGCTT ATGCCATTAT TTCTGCATAT
5101 CAGAGACTTT TAGTACTTTC ACTAATTTTG CTACTGTTAT CTAAAAGGCA
5151 GAGCCAGGTG TACAGGGTGG AACATGAAAG TGGACTAGGA GCGTGACCAC
5201 TGAAGCACAG CATCACAGGG AGATGGTTAG GCCTCCGGAT AACTGCGGGT
5251 GGGCCTGACT GATGTCAGGC CGTCCCACAA GAGGTGGAGG AGTAGAGTCT
5301 TCTCTAAACT CCCCCGGGGA AAGGGAGATT CCCTTTCCCG GTATGCTAAG
5351 TAGCGGGTGT TTTTCCTTGA CACTGACGCT ACCGCTAGAC CACGGTTGGG
5401 TCCGCTTGGC AACGGGCCTC TTCCCAGATG CTGGCGTTAC CGCTAGACCA
5451 AGGAGCCCTC TAGTGGCCTT GTCCGGGCTT AACAGAAGGC TCTCACTCTT
5501 GTCTTCTGGT CACTTCTCAC TATGTCTCTT CAGCTCCTAT CTCTGTATGG
5551 CCTGGTTTTT CCTAGGTTAT GATTGTAGAG CGAGGATTAT TATAATATTG
5601 GAATAAAGAG TAATTGCTAC AAACTAATGA TTAATGATAT TCATATATAA
5651 TCATATGTAT GATCTAGATC TAGTATAACT CTTGTTGTTT TATATATTTT
5701 ATTATACTGG AACAGCTCGT GCCCTCGGTC TCTTGCCTTG GCACCAAGGT
5751 GGCTTGCCAC CCACAGCCTC TCGAGTAGCT GGGATTACAG CCATGTGCCA
5801 CCATGCCTGG CTAATTTTTG TATTTTTGGT AGAGACAGGT TTTCACCTTG
5851 TTGGTCAGGC TGGTCTCGAA CTCCTGACCT CGTGATCCCC CACCCCCCAC
5901 CCCCAGCCTC CCAAAGTGCT GGGATTACAG GCGTGAGCCA CTGCACCTGG
5951 CTGAGTTGGA GCTTTTCTTC CCTCTTTTTG GACTTTGGAA AATGCTCTTG
6001 GTCCATGATG CTATGTAGAC AGCTCCCGTT GACTGTGGCC TGTGCGGCAT
6051 TGGGCAGCAC TCTGGTGAAC ACTGAATCGG GTCTGACCTC CTAGCCCCAC
6101 CATTTACTGG CTGAGCCTCA GTTTCCTTGC CTGTAAAATC AGGAAGATGC
6151 TGGCTCTGCT CCTCTCTGCA CATTTCCCCG TCCTAACAAC ATTATAACTG
6201 TTAGGAAAGA GACGGGCTTG TTTTGGGATG GCTCATTTTA TGTGACCCTG
6251 TGCGCTGTCT CTGAGTCCAT CTGCCCTTCT TCCAGGGTGT AGGGACCAGC
```

FIGURE 3B

```
6301  CCCACAGGGT  CGGTGGGTCT  CTCCCTGTGT  GCGGCGATGA  GAGAGTGTAG
6351  AAATAAAGAC  ACAAGACAAA  GAGATAAAAG  ACAGCTGGGC  CCGGGGGACC
6401  ACTGCCACCA  ATGCATGGAG  ACCAGTAGTG  GCCCCGAATG  TCTGGCTGTG
6451  CTGTTATTTA  TTGGATACAA  AGCAAAAGGG  GCAGGGTAAA  GAGTGTGAGT
6501  CATCTCCAGT  GATAGGTAAG  GTCACATGGG  TCACGTGTCC  ACTGGGACAG
6551  GGGGCCCTTC  CCTGCCTGGC  AGCCGAGGCA  GAGAGAGGAG  ACACAGAGAA
6601  AGAAAACTTA  TGCCATTATT  TCTGCATATC  AGAGACTTTT  AGTACTTTCA
6651  CTAATTGACT  ACTGCTATCT  AGAAGGCAGA  GCCAGGTGTA  CAGGATGGAA
6701  CATGAAGGCG  GACTAGGAGC  GTGACCACTG  AAGCACAGCA  TCACAGGGAG
6751  ACAGGCCTCC  GGATAACTGC  GGGCAGGTCT  GACTAATGTG  AGGCCCTCCA
6801  CAAGAGGTGG  AGGAGCAGAG  TCTTCTCTAA  ATTCCCCCGG  GGAAAGGGAG
6851  CCTCCCTTTC  CCGGTCTGCT  AAGTAGCGGG  TGTTGTTCCT  TGACACTTTT
6901  CGCTACCGCT  AGACCACCGT  CCGCTCGGCA  ACGGGCGTCT  TCCCAGACGC
6951  TGGCGTTACC  ACTAGACCAA  GGAGCCCTTT  TGCTGGCCCC  GTCCGGGCAT
7001  AACAGAAGGC  TCGCACTCCT  GTCTTCTGGT  CACACCTCAC  TATGTCCCCT
7051  CAGCTCCTAT  CTCTGTATGG  CCTGGTTTTT  CCTAGGTTAT  GATTGTAGAG
7101  CGAGGATTAT  TATAATATTG  GGATAAAGAG  TAATTACTAC  AAACTAATGA
7151  TTAATGATAT  TCATATATCT  CTAAGATCTA  TATCTGGTAT  AACTATTCTT
7201  GTTTTATATT  TTATTATACT  GGAACAGCTC  GTGTCCTCGG  TCTCTTGCCT
7251  TGGCGCCTGG  GTGGCTTGCC  GCCCACACAG  GGCATGTCTG  GATGGTTTGA
7301  ACACTAGGGC  TTCTGATGCT  CTAAGCCAGA  GTCAGGTATT  CATTCCATGG
7351  CACATGTGGC  TGGGGTCTGC  CCTGAGACCT  GTCCCGTGCC  AGGCTCTGGG
7401  GGCACATGGC  TGATGGAACC  AAGCATGGGG  AGTGAAGGTG  GAGGGTGGCC
7451  TGTGAGCACC  ATGCCTGAGA  GGACCAGGCT  GGGGACGGAA  GGTTCTTAGT
7501  GGATAATATT  TATTGTCTCT  GCCTCCCCCC  TGACATTTGC  AAAGCGGCAT
7551  ATGCTTGTAA  AAAAATTTTG  AAACAGAAAA  ATATAAATAA  ATAAGTAGGT
7601  ATTACCACAT  GCAAGGGTGA  CCAATTTTGT  ATTTTTCTTC  CCAGCAGATG
7651  TTAAAGCAAG  ACCAACAGTC  TCCCCTCATG  GAAGGCCCAC  TGATCTAAAA
7701  TGCTGGTTCC  TTTTGGACCT  TCAGGGCACT  TGGGGGAGAC  CTTCCTGAGG
7751  TGCTGTGCAG  TGTCTGGTGT  TTCTCAGACC  CAGGTGGTCA  TGGGAGCCAG
7801  GCGTGGCTGA  GTGGGCTCTA  CAGGCCCTAG  GCAGGGAGCA  TCGCCTGTGC
7851  TGTGGCTGAC  GTTCCTTCTG  GCCCTGTTCC  CAAAGTTCCC  CATGGGGGCC
7901  TGGGAGGAAT  GGCCTTTCCA  GGGGGTGTTT  TTATGAGAAG  GAGGTAGCTC
7951  CCTGTTGGAG  TGAGGTGCTC  AGGAGGAAAG  GGGCCTGGTC  TTAGCAGTCA
8001  TGACCACCTG  TCCCCAGTGA  GGAACATCTC  TCCTGCCACA  CAGGCCTCCT
8051  GATCACTTGC  CACGTGGCAC  GCATCCCTCT  GCCAGCCGGA  TACAGAGAAG
8101  AGATTGTGCG  GTACCTGCGG  TCAGTGCAGC  TCCCTGACGG  TGGCTGGGGC
8151  CTGTGAGTGT  GCCTGCCCCT  GTGTCACTGC  ACATGTGCAT  GTGTGTGTTC
8201  TCATGATGTA  GGAGATGCTT  GGGTTTCCAG  GCAGCTGCCA  GGGGTTAGGA
8251  GTGATTGCAG  CTGTGGGTGT  GGGGTGGGTG  AGGGAGAGAC  TAGCAGGCGG
8301  GGAGTGGGCT  GAAGGCCATG  CAGGTGGGGC  CTCGGCTTCA  CATCTTTTGT
8351  TAAATGGATT  TTGTGGCTGT  TACGACACTC  TTGAGACCCA  CATGTGAAAA
8401  CTGTCAGTCT  GTTATCACTT  AAGACAGAAG  AAAATTGCCC  TTGACTCTGG
8451  GCTGGCAGCA  GGTGGAGACA  AGGCCTGACA  GCTTTCCTGC  CATGTGGCAC
8501  ACACTTTGGG  AGCAGAGCCA  TAGCCCAAAG  TGGACCGCCC  TTGAGCTAGA
8551  AGTGTTGACT  CAGGCGTGGG  AAGGTGTAGA  GCAGGCGGGT  CACGGTGAGG
8601  AAGGAGTGGG  GGGCTCAGTT  GTCATGGGAG  GTGCATGAAT  TCGTACTGCA
8651  GAGTGGCTGC  TCAGGGGTCT  CCTGTGTTGA  CATGTTATGT  CAGGTTAAGC
8701  CATTTTAGCA  TTCTTAGTTT  TCTGAGGAAA  CTCCACAGAA  AGTTTTGCTT
8751  TATTTCTTAG  AAGTAAGGAC  AGATACCGGT  TTCTCACCTG  TCCTCTGCTC
8801  CTGTAGGCAC  ATTGAGGATA  AGTCCACCGT  GTTTGGGACT  GCGCTCAACT
8851  ATGTGTCTCT  CAGAATTCTG  GGTGTTGGGC  CTGACGATCC  TGACCTGGTA
8901  CGAGCCCGGA  ACATTCTTCA  CAAGAAAGGT  ACGGCATGTG  CAGCATGTGC
8951  TGGGCCAGGG  GTTCGTGTCA  ACTCGATAAT  GAGCTCTCAC  AAACGAGATA
9001  CAGAAGATG   CACTTGCAGC  TGAAACAGTG  GGCAAAGCA   CATGAGCAGG
9051  GAATTTGTCA  AAGCAGAAGT  AGGCAGACAC  TGTTTAACCT  AGGCATCATT
9101  TTTTAAAAAA  GCAAATTAAG  AGCCAGGCAC  AGTGAGTGGC  TCACGCCTGC
9151  AATTCCAGCA  CTTTGGGAGA  CTGAGGTAGA  AGGACCACTT  CAACCTAAGA
9201  GTTCGAGGCC  AGCCTGGGCA  ACATAGTGAG  ACCTGGTCTC  TACAAAAACA
9251  ATAAAATATT  AGCCAGGTGT  GATGATATGC  ACCTGTAGTC  TCAGCTACTT
9301  GGAGGCTAGT  AAGGCAGGAG  GATCACTTGA  GCCCAGGAGT  TCTGGGTTGC
9351  AATGAGCTGG  TTGTACTACT  GCACTCTAGC  CTGGGTGACA  GAGTGCGACC
9401  CTGTCTCTAA  TAAAATAAAA  AAGCCAAGCA  AACTAAGACA  ACCAGGTAAT
```

FIGURE 3C

```
 9451 TCTGTTTGTT TCCTGAATTG GCAAAAACTT AAACGAACCG TGTTAATATG
 9501 TCCACCTTCT GGGGGGCAGC CTGGCTGCAG GCAAGAGCAG CCCTGGAGCT
 9551 TGCACCTTCC AAGCTGATCG TCTACCTCTC CAAGCCCGGG GCTGTCCACC
 9601 TCTCCAAGCC CGGGGCTGTC CACCTCTCCA AGCCCGGGGC TGTCCACCTC
 9651 TCCAAGCCCC GGGCTGTCCA CCTCTCCAAG CCCCGGGTTG TCTTACCTCT
 9701 CCAAGCCCCA ACTGTCTACT TCTCCAAGCC CTGGTCTGGC TACCTCTCCA
 9751 AGCCCTGGGC TGTCCACCTC TCCAAGCGCC AACTATCTTT CTCTCCAAGC
 9801 CCTGGCCTGG CTACCTCTCC AAGCCCCAGG CTGTCCACCT CTCCAAGCCC
 9851 CAACTGTCTA CCTCTCCAAG CCCCGGCCTG GCTACCTCTC CAAGCCCCTG
 9901 GCTACCTCTC CATGCCCGGC CTGGCTACCT CTCCTCTTGC CTATAGGCCC
 9951 TGAGGGGCAA TTCCAGCCCA AGGGAATCCA TGGCTCCTGC TGCTCCAAGA
10001 AAACCTAGTT TATGTTGTGG CTCTGCAGAG CCTGGCCTGG TCTTGTCCTC
10051 TGTGTTTCAC AGACCTTCCG TAGCCAGTCC CACCTGCCCT GCTCTCTGCT
10101 GCATGCGCAG GGGCCTCCTG TCAGCTCCTC AGAGACCCTT ATTATCCCAG
10151 GGCTCGCCAT GCACTGCCTC CTTCGCCTGG AGCCTCTTAC CTTCCACTCC
10201 TGCCCCGCTG GCTCACACTT TACGTGTTCC TTCTTTGAGG ACCTCTTCCT
10251 GACCTACCGT GCCAGGTGGA GTGTCCTGTT ACGCATTCTC ATGAGATCCT
10301 GCCTTCTTTC TTGGTGAGCT TGTCACTATT GTCCTCAGTT CACTGTCAGC
10351 CTTTGGTGTC GTTGATGCTG CGTCCCCAAG GCTGCTGTCC GGTTCCCACC
10401 ACACTCCTGG CGCCTGCCTG GTGAAGGAAC GTGTTTAGGC TGCACTTTGC
10451 CTAGTAGCTT TGTGGGTCTT TATTGACTTT TGCATACCTT TTGGGGTTTG
10501 GAGCAGGGAC TCCTCAGAAG CATGTTTAGA TGGTGTGGCT GTGCCAGGAC
10551 TGCTGCTGCT GAAGTGGCTC TGGCATGGGG CCAGCGTGCT GGAGCTACTC
10601 TGGAGTCTAG GGTCGTCTTT GTTCCCATAC AGGACCAGTC TGCCAAGTGG
10651 AGATGACACA GACTGGGGCA GCTCAGGCTT GGCTCAGAGG CGCGAGGCTGA
10701 GTGTGCGCTG TCACTTCCCC ACCTTGCCTT CTCCAGGCGC ATGTGCACCT
10751 GGGCCCCTCG CTCACCTGAG CACTGAGGTG TCCCTGGACC TTCCCAGGTA
10801 GCTGTCTTCA TGTGCTCCTT CCTGGGGCCA GGGGTTGCAA ACACCTCTCC
10851 TGGGGCTGGA CACACACACT CCCAGGAAAG CCACTGGTTC CACCTAGGGG
10901 GCCGTGTATC CAGGCAAGTT CTCAGCACTC TGGAACCTGC TTCGCACATG
10951 GGGGTCGCAA GATCCACATG AGGCTGCCCT TGCCTCATGG AGAGGGGCAC
11001 ACGTGACTCC CAGAGGGTGA AGCTTCCCAG CTAGAGGCAG TGCAGACTTT
11051 GCTGACAGGA AGCAGATGAC GTGGGCCTAT TCTCTCCCCG CTCAGGTGGT
11101 GCTGTGGCCA TCCCCTCCTG GGGGAAGTTC TGGCTGGCTG TCCTGAATGT
11151 TTACAGCTGG GAAGGCCTCA ATACCCTGTT CCCAGAGATG TGGTATGTCT
11201 GCTGTTGATT GGGTTGTTGG GTCGCTGCTG CTGTCCCGGG GAGTAGAGTG
11251 ACAGGGACCG TGGGTCAGGT GCAGGCTGTG ACAGCAGAGA GGGGTGGGCA
11301 TTCTGTGGGT GGGTGGAGTT AGGCTCCTGG CAGAGGCCCT GATCAAGCTT
11351 GAGTCCTGTA GGGGTACAGA AAGGGGGAGG TTCCCAATTG AGCAGGAAGA
11401 AGGCTGTGCC ATGGATGGAG GTACCCCGAG TCAGGCTGCA GGCAGGGCTG
11451 GGTGGCTTCC CTCTTGCTGT GGAAGACTCA GCATCTGTAG AAGTGGGGGG
11501 GTGCCCCTCC CCCAGCCTGC ACAGGGGCGT CCTGTGTTGC TGCTGCTGCG
11551 TTTGTCTCCT TTGCTGGTGA ATGTGAAGTG TGTCCCGACG TGACACCTCA
11601 CCTGTGGACT CAGCGTGTGT GCCTTTAAAA GATCAGTGTC TGTGGCCAGG
11651 TGGGGTGGCT CATGCCTGTA ATCCCAGCAC TTCGGGAGGC CGAGGCGGGC
11701 AGATCACGAG GTCAAGGGAT CGAGACCATC CTGGCCAACA TAGTGAAATC
11751 CCGTCTCTAC TAAAAATACA AAAATTAGCT GGGCGTGGCG GCGCGTGCCT
11801 CTAGTTTCCA GCTACTCGGG AGGCTGAGGC AGGAGAATCA CTTGACCCTG
11851 GGAGGCAGAG GTTACCGTGA GCCGAGATCG TGCCACCATA TTCCAGCCTG
11901 GCGACGGAGT GAGACTCTGT CTCAAAAAAA AAAAAAAAA GATCAGTGTT
11951 TGTTTTTTTA AACAGAACCA CATACTGTTT AAATACCCAG CAAAATCAAC
12001 ATTAATTTCT TATTATCTGG TGTGTGTTTT TTTTGTTTTG TTTTGAGACG
12051 GAGTCTAGCT TTGTCACTCA GGCTGGAGTG CAGTGGCGTG ATTTGGGGTC
12101 ACTGCAACCT CCGCCTCCCG GATTCAAGCA ATTCTCCTGC CTCTGCCTCC
12151 CGAGTAGCTG GGATTACAGT CTCAGGCCAT CACGCCCAGC TAATTGTTGT
12201 ATTTTTAGTA GAGACAGGGT TTCACTATGT TGGCCAGGAT GGTCTCAAAC
12251 TCCTGACCTC AGGTGATCCG CCTGCCTTGG CCTCCCAAAA GTGCTGGGAG
12301 CCATGAGCCA CTGCTCCCGG CCTTATGTGG TGTCTTTAAC CAGTGTCTTG
12351 TAACATTTTA TGGCTATCTA TTGAAAGCAG TGGACATCTC CCCAGAAAAC
12401 ACTCGTGCAT ATGAGTTTAC CCCGTTATGC ATTTTGGGAA GTGAGACCCT
12451 GGAACCACAC AGAGCCCCTG CTGGCTTCCT TGAGTGTTGT GGGAACCCTG
12501 GTGGGGGTGT CCCCTACAGA GCTATCATCA GGGCTGGGGG GGTCCCTTGT
12551 GTTAGATGAC TTTGGTGCGG GGGTGGGGGG TGGGGGGTCA AGTTAGGGGA
```

FIGURE 3D

```
12601 GGCAGGAAGT GAAGGGGCCG CTCAAGAAAG GACAGCAGCA GTGTCCTGAT
12651 GCAAAGGCCG GGGGCTTAAC CCCGGAAGCC AGTTTGGGTG GTGACGGGGA
12701 GGCACAGGGA TGGTGAGATC ACCCCGGGAG GGTAGACAGA GATACCAGAG
12751 TAGGGGGCAG GGTTAGGGTG CCGCTACCTG AGGCGGGCCG TAGAGCACAT
12801 AGGTTGGGAG GTGTCCTGGG GCCATTCAAA TGCCCGCTGG ACTCTGCGCC
12851 TCGCCCGTGT GTAATGAGCG GCAGAGGAAG GACTGAGACG GCAGTCAGCA
12901 CAGCTGCCAG GGCAGGAGGG GTGTGGGTTC CACACGCTGG TGCTGGTGAG
12951 GGCGTCTCAT CTGCCCCACT TGGGGGGGCC GTCGGTCAGT GCTGCCGCAT
13001 GGGCACGCCA GGGTGCTGCT TGTCTTTGCT GGAGTTGCTT GGAGGGTGGG
13051 TTGGGAGGTG AAAGGAGGAC CACAGACCTG AACCACTCCA GCTGCGAAAT
13101 GCTGGAAGTG TAACCCAAAA TGTGAGAAAA AAAAACACCC TTTTAAGTAA
13151 GTGGGTGTGA AAGTGGGCCA AGGCCTGATG CCACAGTCAG GGAGCAGGGA
13201 AGGCTCAGCA TTGCTCACCC TCACTTAAGG ATGGGGCTAG CATCACATAA
13251 GGCATCACAT AAGGATGGGG GCTAGCAGGG AAAGGGAGAG AAAACACATG
13301 AGGCACACAC AGACCCTGGG AAGCTGGTGG AGCTGTGCTA ACGTCAGCAG
13351 ACCAGTGATC AAAGACCCAG GCCTTGGGGA GATTCCACAG ACCTACAGAC
13401 CTACAGTTTC TTTTTTCTTT CTTTTTCTTT TTTTTTTTTT TTTTGAGACA
13451 GAGTCTCTCG CTCTGTCACC AGGCTGTGTG CAGTGGCACA ATCTCGGCTC
13501 ACTGCAACCT CCTCCCAGGT TCAAGCGATT CTCCTGCCTC AGCCTCCCGA
13551 GTAGCTGGGA CTAGAGGCAC ACACCACCAT GCCTGGCTTA TTTTTGTATT
13601 TTTAGTAGAG ATGGGGTTTC GCCATGTTGG TCAGGCTGGT CTCAAACTCC
13651 TGACCTCAAG TGATCCACCA GCCTCGGCCT CCCAAAGTGC TAGGGTTACA
13701 GGCGTGAGCC ACCGTGCCCC TCCTAAAGTT TCTTAAATAC ACTTTTAAAA
13751 GTAAACTTTA AATTTTGGAG TAGTTTCCAA TTTCTGGAAA AGTTGCAAAG
13801 ATAGCCAAGA GTGTTCCCTG GGGCCCTCAC ACCATATCCC CATTGTTGAT
13851 GTTTTATGTT ACCAAGGTAC GTTTGTTGTA GCTAAGAAAC CCACGGACAA
13901 TCCTAAGCAT TTAGGAGCTC CATCACCTGG TTTTAGGATG CAAAATGCTG
13951 ACCGAACTAG GAGGTGCAGC TCCTCAGAGG GTGCACCTAT GGTTCAACTG
14001 TGCCCCTCAG GAGCACGGTT GGGAAATGCC CGCAGATGCA CTGACGTGGT
14051 GGGGAATAGC CATCCACCAG TGTTCGTGCT TGAAAGGGCC CAAGGTATGG
14101 ATGCTGGCGG AGGGGGCAGG CTTGAGTCTT GGGGTCTCCC ACTGACTCCT
14151 GCTGTTGCCC CAGGCTGTTT CCTGACTGGG CACCGGCACA CCCCTCCACA
14201 CTCTGGTGCC ACTGCCGGCA GGTGTACCTG CCCATGAGCT ACTGCTACGC
14251 CGTTCGGCTG AGTGCCGCGG AAGACCCGCT GGTCCAGAGC CTCCGCCAGG
14301 TAGGACCTCA TCAGGGAACA AAGTGAAGGC CTCTGGGGCT GGGACCCACA
14351 GGGCCTGGGG CTTCTGGAAT CTAACCACAC CTGTCCACTC ACCTGGTGGC
14401 CCTGTGGAGC GGAGAGCCCT GTGGAGCAGA GCCTCCACCT TCCTCCATCC
14451 TATAATAAAC AGTGAGCAAG CTCTGCCCAG AGGGGACTTG TGCTATGGGA
14501 CAGTCAGTAG CTGTAGCCCA GGGTTCCTGG GGGGGACTTC CAGGACTCAA
14551 GGGATGCAGG AGGCAGATGT GCACTGTGTC CTCTGGAAGC AGGCCTGAGG
14601 CGAGGTTTGA GGTGCAGGAT GTTTATCAGG CCTGCCATGG GGAAGAAGGA
14651 GGGGCAGAGG GAGGAAATGA GCTTCTGGGC AGACCTGGGA CTCATGGAGC
14701 TGGGGAGCTC CTCAGAGCGG TCCTCCCATA GGGGGCCTTC ATGTGCCCTC
14751 GGGGTCAGTT GCTGGAGGGA CCCCCACCCA GGAAGGGACT GGCCCAGGGC
14801 CCTGAGGGCG GATGGTGGGA GGCCACCCCT CCTGGTTTGA GCCAGGCCTA
14851 CCAGGTGCTC CCAGGCCCCA AGGCTCAGAC ACTGCCCCTA CCAGGAGCTC
14901 TATGTGGAGG ACTTCGCCAG CATTGACTGG CTGGCGCAGA GGAACAACGT
14951 GGCCCCCGAC GAGCTGTACA CGCCGCACAG CTGGCTGCTC CGCGTGGTAT
15001 ATGGTGAGCG CCTCCTGAGG GGCCGGCAGG GCAGCCCAGG GTCAGGGTCA
15051 GGGTGTCGCC CACTCATTCA CGCACTCATC CCCTGCCAGC GGCACTGGGC
15101 CACCTCCTCT GTGCCAGGCC CCAGGGGGCG GGATCTCATC GCCCTGCCCC
15151 TCCACCCTGA GAACCAGCTG GTCTTCTACT CTCAGGAGTC CACCCTGTGC
15201 AAGGGTGTGT GGTAGGAGGT GTGGGGCAGC CCCTCCTGGG CAGGGAAGGA
15251 GGAGCTCAGA GACCAGGCCT GGGGGTGGGT GGGAGGGGGA AACCCTGGGG
15301 AAGGGCAAGT CCAGGCGTTG CAGTGCATGG AGCTCCAGGC TGAGGCCAGT
15351 GTCATGGTGT CTGGCATCCA CTGACCCCTG TCCCCTGTAG CGCTCCTCAA
15401 CCTGTATGAG CACCACCACA GTGCCCACCT GCGGCAGCGG GCCGTGCAGA
15451 AGCTGTATGA ACACATTGTG GCCGACGACC GATTCACCAA GAGCATCAGC
15501 ATCGGCCCGG TCAGTGCCCC TGCCCGGCCT CTGACTGCAG CCCCTGGGGG
15551 TTGAGGTCCG AAAGTGAAGT CCTAGAGGCC GGGCTGTGAG CTGGGAGTGG
15601 GGTTTCCTGG AGCCTGGTGT ACCTCCATTT GGGAGGTGGC CCTCTGATCG
15651 CACAAGTGTC TGAGGGCTTC TGTCCTGGAC CCCTGCACGC CCAGCTCAGT
15701 GAAGTTGCCC CACCACACTC GACCCCCCGC TTCCGTCCCC CACCGGCTCT
```

FIGURE 3E

```
15751 TGTCCTCAGT GTGCCTGGAC ACTCTCCTAG AGGCCCCTCC CTGAGATCTT
15801 GCTGGCTAGC TGGCTAGCTG GGAGGGGTGC TTTTTCCTCA CTTGGTTCCC
15851 TCTCCCCAAA CAGTTCATCA TTCGCCATTC TCCCGTGGGG TTTAGACATG
15901 CCCAGGGTGG GTGGGAGTAG CAGGTGCCAC TCCTGATTCC TCCTGCCTAG
15951 CTAGGGACTT GGAGCTCTCA CCTCTGTGGG GCCTGCAGGG GTCCAGGTGT
16001 GGCCAGTTCA GTGACCTTAG AGGGTGCAAT CCCCGGGCTG TGCTGGTGCG
16051 TGGCCGCCTC CTGACAGAGT CAGCAGGCCC TGGGCTGTGC TGCAGCTGCT
16101 GCCGTAGCTG TGCGCGTAGC TGCTGCGGTG TAGTGGGTTG GCTTAGGCAT
16151 TCTCTGGACA TACCCAGGTG GCACTGGGCC ACTGAGTCCC ACCCTGACAC
16201 TGCATCTCGG ATTTTCTGGG CCTCATGCCA CCTCAGTGGA TCACAAATCC
16251 TGACTGACCC TGCAGCGGGT CCCTTGTTTT TTGCTCAGCA GTGATGTGGT
16301 TCTTTGTGGG TTTTGGTTTA ATCCCATATA GAGCACATCT GTACTAAACG
16351 CATTAGAAAC ATGCTTGCAA TTGGATCTTG ACTTGTGAGA TGCATAAGTA
16401 AAAAGTTGGG GGCCTCTGGA ACATTCTGTT CTGAGGAAGA AGGGGGGCAA
16451 GTGGTCCCTA CTGCTACAGT CCTGTCTTCG CATCTCTTCC TGGGCCCCTC
16501 AGGCCCTGTC CTCTGTCCCC TGTGTTGTCT CTAAGGCACC TGGTAGCCCA
16551 TGCCCCTCTG GTTTCTCCTG GAACCCCTCG CTTCTCCCTG GTGGAGTGCT
16601 GCTCCTTCTC ACAGCCTAAG GCAGGCTGTG GCCTTGGCCG ACACTGCCTC
16651 TGTCTGAGTT GGGTCCTGGG GACACAGTTG TTGCCCATCC TCGCTCAGGA
16701 AATGCCTGTT AGAGCAGAAG GCCCCTGTCC TGGCCCTGAG TGATCTGCAC
16751 GGCACTTTAT GCCTGGGGGC TGCTGTGGAT CTGGACGAGA CCTTGTCCCT
16801 GGAGGCTGCT GTGGGTCTGG AGCGGAGCCT TGACAGGGCT GTCTCTCCTG
16851 CAGATCTCGA AAACCATCAA CATGCTTGTG CGCTGGTATG TGGACGGGCC
16901 CGCCTCCACT GCCTTCCAGG AGCATGTCTC CAGAATCCCG GACTATCTCT
16951 GGTGAGTGTG GCTGGGATAT GCTGGCGGGG CCTCTCACGA AGACTGGATC
17001 TGAGCCCCAG CTGCATCCCA GTGAGGGGGC CCCCACGGTG CCATCTGGGA
17051 ATACTGCCAG GGAATACCTC CAGGAACCAG CAGTGTCAGG GCTTGTGGAA
17101 GCCACTGAGG GTTGTCTTTG AATTGGAAGA TTTGCCACCC AGTGGAAGTG
17151 TGGGGTGTTC CCAGAAGGTA GAGTGAGGAA GGGGGTGGTA GGTAGCAGGG
17201 CAGGTTCAGG TTGGCATCAG GAGGCCTGTG GACAAGGGGA GCTTGTCAGC
17251 CATGGACTGT GCCCTGGAGG TGGGGCCCCT GTCATGGAGG GCAGAGAGCC
17301 GTCCCATGGT GGGAAGCTTC CGCTGTACAG GCCTCTTCCT CTGGTGCCTC
17351 AGCACTGCAC GAGGGCGGCA GGGCTGGCAC AGCCTGGGGT CGGGGAGCCT
17401 CCCGCTGCCC CTTGCCTTGG GTGTGGCCCT TCTGGGTGAG TGTGTCCTGT
17451 TTTCCATAGA GTGTGGCCCT CACCCCCAGG AGCCCAGCAG CCCAGCTGGG
17501 GTGGCATCCA GGCCAGTGCC AGGCCTCGGG AGGGGACAGA CGGCCTCTCT
17551 GGGACCCTCC TGAGTGCAGG GTCTGGGTAG CAGCTGGGCT TCCAGCTTTC
17601 TCCTTGCACC TGACTTGGGC TTTTTTCTCC TCACAGGATG GGCCTTGACG
17651 GCATGAAAAT GCAGGTAAGG GCTGCGGGAC TGCGGCTGCA TGCTTCCTTT
17701 GCAATCATGT CTCCCCTTTA TTATTTTTCC TTTGGGGTTC AGAAATAACT
17751 CCTCCTGGAC CAGGTCCCGG CAGCGTGCGA CTAGAGGCTG AGTCAGTTGA
17801 GGCCTCTGGC CGTGTCCCTG TGGGTGCTGT TGGTCTCTGT GTGGGTGCCC
17851 ACCGTTCTCG ATGTCTGTCT GCAGCTGTCC TGTTTGCTTT TTGCCCTGAT
17901 GATCTGAGTG GGCTCAGCTG TGTAACGACA GACCCAGAGC TGCAGAAGCT
17951 CTCATCTTGT TACTGTGGCA GGAGGTGGCT CTGGTTAGTG GGGGCTTCTC
18001 CTCCATGCAC TCTTAATTTA AGGGGCTTCT TCTTAAAGGT CCTGGGTGGA
18051 CAGGACAGGA GCCTGGAGGA CCGTGGTGGC GTGTGGCCGG GCCTGGGAGC
18101 TCCCCGTGGA CTTGGCCTGA GTGGGCTGGA ACCCAGTCAT GAGGGGCACC
18151 AAGCACAAGG AGAGGGGAGG CCGGGTGGAT CCTGGCTGAC CCTGGTCCTG
18201 TCCTGGCTCT GGGGGCCCTG TAGACCGCAG TCCTGTCCGA CTGGGCTGAG
18251 CCTGCGCCCC TCTGTGCGTG TCAGAAGCCC AGACAGTGTT GCCCTGTGTC
18301 TTGTGGTCTA AGGAGGGTTA CGCCCTGCGG TGCCTGTCTT CTGTCCCCCA
18351 CCTGATTCAG TGTGGAAATG TGGAGTCTCC AGAGGTGTCC TGGGTGTCAC
18401 ATTTGGGATG GATACACGTG GGCCCAGCAC TGCCCGCCCC AGGGCTACCC
18451 TTGGTGCCAG GTGCCCCCAG CCACGAGCTT TTACCCAGCT GGCCTTGAGC
18501 TCCCCAGAGG CTCCCCGGAC ACTGTCCGTG TTTTGTGAAA AGGTTTTCAA
18551 AACACATGTA AAGTGGAGGT GAGTAGCAAG CCCTAGAGCA GGCCCTGGCC
18601 TCCCTGCCCC TCCCTGTCCC CTCCCTGCCC CTCCCTGCCC AGCGCTCCCT
18651 CAGCACCGAC TCATCAGTGC ACCTCAAGCT GATGAGGGCG TCTGTGTTTT
18701 GACAAAATTG CTCTGAGGTT GTCACACCCA ACAAACTTAT GACGGTTCCT
18751 GAGTGTAGTC CTCACGTTGT GGCTGGTGTT TGTGAATCAG GATTCAGGCC
18801 AGGCCTGCAC AGGCCTTCAG TTGTTGGTCT TTGAGCTCCT GTTAGTCCAG
18851 CCGTCTCTCG TGGTCTCTTT TCTCCTCCTG GAAGGTTTGT TCCTGAAGGG
```

FIGURE 3F

```
18901 CTTCACATTG CAGATCTGAC TGGTTGCTTC TTATGTTCCC TGAGTTTTTG
18951 TAAACTGGCC AGGCCCTGAG GCTCGATCCC ATTGTGTTTC TTTGGCGAGA
19001 ATGCTTTTCT GGTGGTCCCT GCCTTGTCCC TCCAGTGCAC GATGTCTGGA
19051 TGCCTCTGCC ACACACCACC CCCTGCCCAG TCCCCATGTC TGTCTGGTCA
19101 GTGCCCAGCT CTGTCTCACT AGGGTTTGGT CACCGGCCCT TTGAACTGAG
19151 ACCAGGCTGT GTACCTGTGA GCCCAGCTCG GGGTGAGATT TGAGGTGGAG
19201 CCTTCCCAGC CCTGTGCAGA ATTCCCATCA CCTCCAGGTG TACTCAGAAA
19251 TGGGGATCAT TGGCCAGGTG CGGTGGCTCA CGCCTGTAAT CCCTACACTT
19301 TGGGAGGCCA AGGTGGGCGG ATCACAAGGT CAGGAGATAG AGACCATCCT
19351 GGCTAACACG GTGAAACCCC GATGCTACTA AAAAATACAA AAAAAATTAG
19401 CTGGATGTGC TGGCAGGAGC CTGTAATCCC AGCTACTCCG GAGGCTGAGG
19451 CAGGAGAATG GCGTGAACCC AGGAGGCGGA GCTTGCAGCG AGCTGAGATC
19501 ACGCCACTGC ACTCCAGCCT GGGCAACAGA GCGAGACTTC ATCTCAAAAA
19551 AAAAAGAAAT GGGGTCATTT CCAGGCATCA CCATGACTGA GGTGCGCCAC
19601 TGTCATTGGG TGAGAGCAGC TGGATGCTCT ATGTGTAGGT GCTGGAGCCT
19651 CTGAGGGATC GTCCAGTCCT AGAAGTGTCC TCAGAGGGAC ACTGTCCTGC
19701 CTGGTGGCCC ATGAAGAAAG GGAGGGCTCC CTGAGTCTCC CTGACGTGTG
19751 TCTGCCTGCA GGGCTCAGCC TTCTCTGAGG CCCTTGTCAG CCATGAGGGG
19801 TGCCCAGGGC TCAGAGCCTG AGGCTGAGCG TTGGCTGGGT GGGAGCCCCC
19851 ACACCTGGCC CTCAGGCGCC CATTGGATCC TGGAGGCAGT GGCTGGGAGT
19901 GGGAGGGGCT GCATCTGCTG CTGTAACACC ATCCTTTGTG TGTAGGGCAC
19951 CAACGGCTCA CAGATCTGGG ACACCGCATT CGCCATCCAG ·GCTCTGCTTG
20001 AGGTTCGTGG CTCCTTCTCT TTTCTCAGCC TCAGCTGACC TTCCTGTGCA
20051 CGTAAGCCCA CGCATCCACC TGAGGGCAGC ACTGCTGGCC. ACACACTTGC
20101 CACTCCTCGA TACTTCCAGT GACCTGGGCT CTGGCCTCTG GCTTCAGAGG
20151 GTCGTGCTGT GGAGGGGGCG GCCTTGGCCA GCAGCCTTGG GTGTTGGGCT
20201 GGGTCGGGGG CCTTGGGAGG GCAGGGGCTG GAGGCTGTGT GAGAAGGGGA.
20251 GTCTGGTGAA GGCTGTTTCT GAGAGTGCAG GCAGGAGTGG GACTCCAGGC
20301 TCTTCTTAGA ACTGGAACTG CTTGGGCCAG GCACGGTGGC TCACACCTGT
20351 AATCCCAGCA CTTTGGGAGG CCGAGGAGGG TGGATCACGA GGTCAGGAGT
20401 TCAAGACCAG CCTGGCCAAG ATGGTGAAAC CCCGTCTCTA CTAAAAGTAC
20451 ACAAAAATTA GCCAAGCGTG GTGGCGGGCA CCTGTAATCC· CAGCTACTTG
20501 GGAGGCTGAG GCAGAGAATT GCTTGAACCC GGGAAGTGGA GGGTGCAGCG
20551 AGCCGAGATT GTGCCACTGC ACTCCAGCCT GGGTGACAGA GAGAGGCTCC
20601 GTCTCAAAAA AAAAAAAAA AAAAAGAAC TGGAACTGTT TGTTATGGGC
20651 ATTCTCGAGC CAGTACTGGA GAAAAACGAG AGTGGATTTT TATGCCGGTG
20701 GGAATGAGGT AGGTGGGATT CTGAAGGTGT TTCTGGAGAG CCCTGAGGGC
20751 TGGGCCACGC AAAGGGCCTG CCTACACAGG GTGCTGGGAA CCCTCTGGGC
20801 ATGGATGCTG GCCAGGCAGG GGGGTGCTGG CATCCATAAA TGGTCTCCTG
20851 CGCCCTTCCA TCTTCAGTCA TATCTCATGG ACTTTTGCTG TTTTGTCTTT
20901 AAAGGTAAGT GCAGCAGGAG ACCCTGGCAC TCTCTGGAGA TGTCTGCTGG
20951 TTTGATTCTG GTCCCCGGTT GGGGCAGGAT GTGGCCAGGA CCATCGGGAA
21001 ACCAGCGCAG CCATGCTGGC CGTGCAAGGG CAGCTGAGCC TCTCTGTCCT
21051 GCTGTCTCTT CCAGGCGGGC GGGCACCACA GGCCCGAGTT TTCGTCCTGC
21101 CTGCAGAAGG CTCATGAGTT CCTGAGGCTC TCACAGGTGA GGCCGGTGCC
21151 TGGGGCTCTG AGGGGGCTGA AGAGGGGGAT CAGGGCTGGG AGCTCCTGCA
21201 GGCAGAAGTG CCCACCTCAC CTCCACCCTG CCCTATTTCC TGCACTGGTG
21251 TTTCAGGGTC ACCCCCACCC TCCCATCCCC TCCCTAGCCC CTGCTCCATC
21301 CACCGGTCCT CCTCGGGCTG GCCTCACCTG GGGCAGTTCT CTGAGGCCTG
21351 CAGGGTGCTG GGGGTGCTGG CAGTTTCTGC GTCCTGCTCA TGTTGGAGCC
21401 ACTGTGTGCA AGGGCCAGGC ACGGGCAGGG GCTGTGTACC CTGAGCTGCA
21451 CAGCCTACAC GGCACCTCCA TGTCTCTGAA GCACCTTCTG CCCATGGAGG
21501 TGACGCCAGC CTGTGGACTT GCCCTCCTGA GACTGTTTGC AGCAAAAGCC
21551 CCGGTCCCTC CTGCCAGATC AGCTGCCCAC AGACCCTGCC CGAGCCCATA
21601 GTTTGACCTC AGTGTCTCTC ACACGTGCCT GCACCCCAGT CTGCAGCCAC
21651 AGTCATCCCA TACATGCGCC CCAACCTCCC GTGTCTCCCA CACCCTGTCC
21701 CGGCCACGGC CTCAGCCAGT GTCCCTCTGC CTGGAACCGC TGCCCCCCAG
21751 CCCCGTCTCC CTCCCTTCAG CTCTCACTAG GACATTGTTC TGCAGGGCTT
21801 CTGGGTCTTC CTGGCCTCTG TGTGGCCAAG GCTGGCACCC ATCTTGGGCT
21851 CAAGCAGAGG AGGGGCATTG TCCTGCTGTG CCTGGCCCAA TGGCGGCCTG
21901 CTCCTGCTCC TGCCTCCTGC CCAGGACTTG CTCTGGGTGA TGGGGACTTG
21951 GGGAGGCTGA CTGAACCCTA CGGCACTCCA GGCCTCTTCC CTTCTCACTG
22001 AGGTGAGAGA GGCAGCCAGA AGCTGAGGTT GTTCAGGAGG CATTGGGGGC
```

FIGURE 3G

```
22051 GCCTGGCACA GAGCACACCC GCAGAGACCT GGGCCCCCTC CCTGCCTTCT
22101 GGCCGGTGGG GAGATCACAG GGGAGTCAGG TGCTGACTCC CAGTCCCGTC
22151 TGGGCTGGTT TGAGCCCTCG CTGGCCAGTC ACGTTTCCCA GCAGCTGTGG
22201 GTGGTGAGCT AAACAGGTGC AGGCCCTCGC GCGCCTCGCA GCACCAGTGG
22251 TGGCTGTGGC CGGCAGAGTA AGCTCCCAGG CACGTTCTGC CTCTCCAGTC
22301 CTGCCCAGTC TGTCTCAGCG ATGTCCCAGA TGGGGACGTC CCGTGGTGAC
22351 GTGTTCTCTG CTTCCACATT TGCCCTCGAT GCTGCCCAGG TCCCAGATAA
22401 CCCTCCCGAC TACCAGAAGT ACTACCGCCA GATGCGCAAG GTATGCGGGA
22451 GCCAGCCCCA TCCCTGTCCC GTCCCCCAGG GGAGGCCGCC CTCAGCAGGG
22501 TGGGTCCTTC CCTCTGAAGG GGGGGCTCCT CCCTGGGGGA CTCCTCCCTT
22551 GGCGTTTTTG GGTGTCCTGC TGTGGTGGAT GCCTGGCCTA GGGGCTCATG
22601 CTTCATGTTG CTGAGCTGCC TGGCACATGG AGGCACAGTT GGCTTGCACA
22651 CACAGCCGTG CCTCAGAGCA GTTCCAGTGG TCACGGCACA CACAGGCTTC
22701 AGAAGGACAG CCGAAGTGTA GCCAGTGTGT CCGGGGAAGG CAGAGGAAAG
22751 AAGTAGACCT CAGAGCCGGT GTGGGCTGTG ACCACAGGTG CAGACTGTGA
22801 AATTAGGCAT GGACCCAGCT GCTGCTGCCT GTTTACAATG GGGGTGGGGG
22851 GCACCTGGGC CCCATCCTGT CCGTCGTGAG ATCTGCAGGT GTTGAGGGTG
22901 TGAGCTGCAC CCCTGAGGGT CCCTGTGCTG GAAGCTGGAG GTCTGTCTGG
22951 ATGTACCCAG CTTGGGGCCC TGGCTGCACC CACACCTTTG GTGGCTGGGC
23001 CCCTGCCCTG ACCGGGTGCT CTGTGGTGGG GAGGGATGCG TGCGGCTGTG
23051 GGGAGGTTCT GAGAACTGGG GTGTGGACAC CCCCAGCCTG GAGTCATGGC
23101 TTGTGCTCTG CAGGGTGGCT TCTCCTTCAG TACGCTGGAC TGCGGCTGGA
23151 TCGTTTCTGA CTGCACGGCT GAGGCCTTGA AGGCTGTGCT GCTCCTGCAG
23201 GAGAAGTGTC CCCATGTCAC CGAGCACATC CCCAGAGAAC GGCTCTGCGA
23251 TGCTGTGGCT GTGGTAAGGC TGTGGTCCCA GCAGCCCCGT CCATACCTCG
23301 TGTCCTGCAG ATGAGCTGCG TGCTCACTTC CACTCCTGTG GGCTCCAGCC
23351 CAGCACACAG TCCGGCCAGG CCGTAGGAGC TTGTCCTTGG ATGGTGTCTA
23401 TATGTGGAGA ACTGTGAGCT CTGGCTGGAC CCCTAGGGGC CTTGCTGGGC
23451 TGTGTGCACA GGGCCCTGCA CTGCGGAGCT GGTGTCCAGC CCAGCCACCG
23501 ATACTTGGGG GAGCCGGCGT GGCCCCCAAG GTTTCTCTCT GGTGGTTTCC
23551 ACTGGGTGTC TGAAGAGGGA ATTTGTTGGT GTTGGTTTTG GTGCCACATC
23601 CTTTCAGCAC ATCTGGCTTT TGTGTGTGTT TCCCAGTGGA GACCCTGCCC
23651 TTTTCTGGCA GCACAGACTT GGTTTCTAAG TCATGGGCAC GTGTGGGGGC
23701 ATGTTCCCTG GTGGCTGTGC ATGGAGGCCC TGACAGATGA GGTTGCAGCT
23751 GCTGCTTGGG GCACCCGAGG GCTTGGTTAA CGTGGAAATC AGCTCTCCGC
23801 CCCCTGTTCC TGCCCCATCG GTTGTCAGCC CTAGTGTTGC CTCTAGAGAG
23851 TTCCGCTGTG CCCTGGGCGC CTGTGTGTGC TCAGCACATG GGCGAGTTCT
23901 AGGGTGCTCT CTGTGATTTC AGCTGCTGAA CATGAGAAAT CCAGATGGAG
23951 GGTTCGCCAC CTATGAGACC AAGCGTGGGG GGCACTTGCT GGAGCTGCTG
24001 AACCCCTCGG AGGTCTTCGG TGAGTGGTCG GCCAGCACTG CGGCGCGCAA
24051 ACCCGGGGCT GGCTAGCACT GTGGTACACA AACCTGGGGG CCAGCTTTTC
24101 CCCCTTGCCC GAGGCTGCAA GGGCCCAGGT TCACCGGCAG ATCTGTCTGG
24151 AGCCCTCCCT CAGCCCAGGC TGTTCTGCGC TCCTCCATCC CCCGGGGTGG
24201 CAGGATCCTT GTGTTGTGGA TAGGAGGGCA TCAGGTCAGA CCTAGGGGAC
24251 AGTGGAGGGT TCCAGTGAGA TCCACAGCCT GGGCTGGTTC CTGCTCAGTC
24301 CACAGGGCTT GTGTTCTGTG GAGGCTGCTG TGTATCCAGA GCGCCTGCAG
24351 GGAGGTGTCT TTGGGGACTG TGGGGACTGT GGGGACCCAT GCCATGGGCA
24401 GTAGGCTGCT GTGTGTGCAT GGTTGCCACC GTACTGGTCT TGGGGGAGGA
24451 TCTCAGCCCT GGTCCACCTC TGGGCACCTC ACATACCCGC CTTCCTGGTC
24501 CCCTCCACAT CACACATGGC TTTTTGGGGT GGGGTCGCAG CTTTCTGCTG
24551 TGTTCCCCTC ATCTTCGCTC TCAGGTAGCA CAGGTGTGTG TCCTGGACCA
24601 GCCGGCGTTT GCTCTGGAGG TTGGTCAGGG AGGCAGCGTC CGGGCCCGGG
24651 CTCACTGCAA CACTCTTGCT TGTTGTGGCT TTGCCTGAGC TGCAGAGCCT
24701 GGGCAGCCAG GGTGAAACCC AACACTTGGT TCTTCCCTCC CTTTCCCAGG
24751 GGACATCATG ATTGACTACA CCTATGTGGA GTGCACCTCA GCCGTGATGC
24801 AGGCGCTTAA GTATTTCCAC AAGCGTTTCC CGGAGCACAG GGCAGCGGAG
24851 ATCCGGTAAG GAGGGTCTCA GCCATTCAGT GTGGGCGCTG CCAAGTCGGG
24901 GGCCAAGACC CAGACGCATC ATTCTGTGAC ACGGCCCTGG TGGCCCATCT
24951 CAGAAGCGAA ACTCATGGAA ACATGCAAGA GGCTTCGGAT GTTGTGGAAT
25001 CCAGTCATAT GCCCTAAAGC ATACAAAATA TCTGTTAGGG GCTCAGAATA
25051 GCACAGTTAT GATACAAAAA TGGATTTTCT CTCTCTTTTA ATAATGTTAA
25101 GAAGACATCA CATACCTGAC TCCACCGGTG TCCCAGAAAC GGTTTTTAAG
25151 TAACCTTTCC TGTTGAAGGG TAGCAAGTAT TCAGAAAAGT GTACAGGTTG
```

FIGURE 3H

```
25201 GTCTTCTTGA AGCAAACAGG AAGCGAACAG TGCCAGCATT AGACATGGTG
25251 ACACCACCAG AGCCCTCGGC CCGCCCCATG ACGGGGCCGC CCACATGCCT
25301 GCCAGGTCGT GGGTGTCTGT TGCTCGCTTT GGATCTTGTC TAGGTGGACT
25351 CCTGAGGTGT GGAATTCGTG TTGCCTTCTC CTGCTCTCCT GCTCTCCTGC
25401 TCGCGGTTAG TCAGGTGGCT CGGGTAACAG CAGCGTTCTC TCCCTCGGGC
25451 CTTCGGTTGA ACACAGAATG CCGCGCTATC CAGCTGCCTG TTCTCAGCAC
25501 CTGGGAGGAT TTCAGTCTGG GTTATTATGA AGCATCTACT GTGAACACTC
25551 TTGTACTTAT CTTTTGGGGG CACCTGGGTA CCCATTTCTC ATGGTCACGT
25601 ACCTAGGAGT GGCATTGCTG TGTTAGAGGG TACGTTATAG GGTATGTGAT
25651 TTTTGTAGGT TCTTCTTTAT CCTATCACGA TTACATTTTT TTACTTTTGT
25701 TCAACCTGGT GTAGACTCAC CTTGGTCACA ATGCACTGTC CTTTTTATAT
25751 ATTGCTAGAT TCAATTTGAA GAATATTTTG TTAGGATTTT AGCAACTCTG
25801 GTTACAAGAG ACGCTGGTCT ATAATTTTTT TTTCTTTATA ATGTTTTTGT
25851 CAGGTTTTCC TGTTAAGATG ATGCTGGACT TAGAAAAGCA GTTGGAAAAT
25901 GCTTTTAAAA TACTCTTTGG AAGAATTTAT GTAATATTCA TAATATTTCT
25951 GCCTTAAATG TTTGGGAAAA ATTACCGGAA ATGCCAGTTG GGCCTGGAGA
26001 TTTCTTTGAG GAAAGTTTTT AAATTAGAAG TTCAATTTCT TTCTTTCTTT
26051 CTTTCTTTCT TTTTTTTTTG AGATGGGTTC TAGCTCTCTC ACTCAGGCTG
26101 GAGTGCGGTG TAATTTCTTT AATAGTTTAT AGGACTGAGC AGATTTTCCA
26151 TTTTTGTATC AGTCTGGGGA GTCTTCCCAT TTCCACTCAG CTTTACACTG
26201 ATTCATGCAA AGTTGTTCAG TGTCCTCTTA GATGGCTCTG AGCCCAACGC
26251 TGACATCCTC CTCTTCCTTC TGAGAATCTT ATACTGATCT TTTGAAAAAA
26301 AAAAAATCTT AGTCTTTGAT TCTGTTTTTA AAGAGACTTT ATTTTTGGTT
26351 TCATCAATTT CTATTGTTTG TTATTTTCTT TCTTTCTTAA TTTTTTTGAG
26401 ATGGAGTCTT GCTCTGTTGC TGAGGTTGGG GAGCAGTGGC GTGATCTCAG
26451 TTCACTGCAA CCTCCGTTTC CGGGGTTCAA GCGATTCTCC TGCCTCAGCC
26501 TCCCGAGTAG CTGGGACTAC AGGTGCTGAC CACCATGACT GGCCAATTTT
26551 TTGGTATTTT TATTAGAGAC AGGGTTTTAC CATGTTGTCC AAGCTGGTCT
26601 TGAACTCCTG ACCTCAGGTG ATCCACCTTC CTTGGCCTCC CAGAGTGCTG
26651 GGATTACAGG TGTGAGCCAC CACACTGGCC TTTGCTATTT TCTTTCTCCT
26701 TTATTTTTCT AACTTGAATA CTTAGATATT TGATTTTCAG GCTTTTATTG
26751 AAATATGAAT TTGAGGCTAT AAATGAGTTT TGAGATATCA TTCAGTTAAA
26801 TGTGTGTTCT GGTGCTTGCT GTGGTAGCAC AGATACTAAA AGTGTTTTCT
26851 GTTTCTACTG TTCTTCTCTG GCCCATGAGT TATGTGGGAG TATGCTGCTT
26901 CATTTACAAT CTGAGAATGT TCTGGTGTGG TTTTTTTGGA AGCCGTGGAT
26951 GGAGCAGGGG TTTTCTTGTG CTTCACAGGT GCAGCTAGGA GGGCACTGTG
27001 TCCAGGGTCT TCTGTCGGCC TGGCGTGGCC CTTGGCCATG TGCTGCTCTG
27051 CGGCATGAGG TGGGCGTGAG TTGTCCTCAG CCACATTTAG AGAATTGGCC
27101 TTTTAAAAAA TAGATCATCT TTTAAAAATC ACTGTAATAA AAGTAAAGCA
27151 GGTTCTTTGC AAACAAGACT TGCAAAATAC AGAGAAGCGC AAAGAAGAAG
27201 CTAAGTCGCC CCTCCTCGCC CCTGAAGGAG AATCTGCTGT TGCTGTTTGG
27251 TCTCCACATT TCCATGGCGG CTTGCTGCCC CTTTCACGCC TGGCCCACTT
27301 TGTGCCTGGT GAGGTTTCTA AAAGCCCCAC CCTTGAGCGC GCTCCTCCAG
27351 CACGAGCAGT AATGGCACAG GTGTTGTGTC ATTTTACTCA GTAGCCTCTG
27401 GGTTATTTTT CAGTTTTCCT TGTTGTTTTT TAGCTTTTCC CCATTTTAAC
27451 CTTAACTGGT ATTTTCTTGT TAAATATTTA TTCATGACCA TTATTATTCC
27501 CTAGAGCCAC ATGGCTTGGG GTCCACCTGC CTGGGTCCGC CCCCATCCCT
27551 GCCCCTTCTG GCTGTCTGAC CTGGCCTGGT GACTTCTCTT CTCTGCTCAT
27601 CTCTCTCCCT GCCTGAGTGG GCAAGAGTAC AGCCTCACAG AGTGGTGGGA
27651 TTGTGTGAGA TGCCACAGGG AAGCACATGT CAGTTGTTGT CACTGTGTAG
27701 AACAATGAGT CCCGGATGTG GCCCGCAGGG GAGCAATGGT GACTTAATCG
27751 CGGGCTTCCT CTGCATTTCT TTGGTGACTT CCAAGCTAGA ACATTCTTTT
27801 TTTGTTTATT TGTTTGAAGC AGGGTCTCAC TCTGTTACCT AGGCTGGAGT
27851 GCAGTAGCAA AATCATGGCT CACCACAGTC TCAAACTTCC GGGCTCAAGC
27901 AATCCTCCCA CCTCAGCCTC CTGAGTAGCT GGGACTACAG GTGCATACCA
27951 TCACCTGTGG CTAATTTTTT AAATGTTTTG TATTTTTTAA ATGTTGCTCA
28001 GGCTGGTCTT GAACTGCTGG GCTCAAGCAA TCCTCCCACC TCGGCCTCCC
28051 CAAATGCTGG GATTACAGAG TGAGCCACCA CACCCAGCCA TTTTTAAAAT
28101 TTTCACCAGG AAGTTTTTTC TTTCATTTTT AAGCACAGTA AGTATTTGTG
28151 TATTATGTTA CAGATATTTT CCCCTCAATT TCTTGTTCT TTTTATCTCT
28201 TTAGGGAGTA TGAACATAAG TTTTTAACTT TTAAATGGTT AAATATATTA
28251 GTGTGATTTT TATATTAAGA TTTTATTTTA TTTATTTTTT TTTTTTTTGA
28301 GACGGAGTCT CGCTCTGTCG CCCAGGCTGG AGTGCAGTGG CGGGATCTCG
```

FIGURE 3I

```
28351 GCTCACTGCA AGCTCCGCCT CCCGGGTTCA CGCCATTCTC CTGCCTCAGC
28401 CTCCCAAGTA GCTGGGACTA CAGGCGCCCG CCACTACGCC CGGCTAATTT
28451 TTTGTATTTT TAGTAGAGAC GGGGTTTCAC CGTTTTAGCC AGGATGGTCT
28501 CGATCTCCTG ACCTCGTGAT CCGCCCGCCT CGGCCTCCCA AAGTGCTGGG
28551 ATTACAGGCG TGAGCCACCG CGCCCGGCCT ATATTAAGAT TTTAAACTTG
28601 CCGGGCGCAG TGGCTGACGC CTGTAATCCC AGCACTTTGG GAGGCCGAGG
28651 CGGGTGGATC ACAAGGTCAG GAGATCGAGA CCATCCTGGC TAACACGGTG
28701 AAACCTTGTC TACTAAAAAT ACAAAAATTA GCCGGGCGTG GTGGCGGGCG
28751 CTTGTAGTCC CAGCTACTCG GGAGGCTGAG GCAGGAGAAT GGCGTGAACC
28801 CGGGAGGTGG AGCTTGCAGT GAGCCGAGAT GGTGCCACTG CACTCCAGCC
28851 TAGGCGAGAG TGCAAGACAC CGTCTCAAAA AAAAAAAAAA GATTTTAAAC
28901 TTACCTGGAG AGTTTTTGAG ATACAGTTTG GAGTTGCAAG TTACTTTAAC
28951 ACTATTTATA TGGAATATTC TATTTTACTA GACAGACTTA AATTCTCCCT
29001 TAAATTCACA AATTTATAGA AAAGTTACAA AAATACTGAA AAGTGCTCCT
29051 GTTTACTCTG ACTAGAATTC TTTAGTGGGT GGCACCCTAC CCTGAGGGCT
29101 TCATGACCTG TCCTCCCACA TGATCCAGGC TCTACCCTCA GGGCTTCATG
29151 ACCTGTCCTC CCACATGATC CAGGCTCTAC CCTCAGGGCT TCATGACCTG
29201 TCCTCCCACG TGATCCAGGC TCTACCCTCA GGGCTTCATG ACCTGTCCTC
29251 CCACATGATC CAGGCTCTAC CCTCAGGGCT TCATGACCTG TCCTCCCACG
29301 TGATCCAGGC TCTACCCTCA GGGCTTCATG ACCTGTCCTC CCACGTGATC
29351 CAGGCTCTAC CCTCAGGGCT TCATGACCTG TCCTCCCACG TGATCCAGGC
29401 TCTACCCTCA GGGCTTCATG ACCCTTCCTC CCACATGATC CAGGCTCTAC
29451 CCTGAGGACT TCATGACCTG TCCTCCCACG TGATCCAGGC TCTACCCTCA
29501 GGGCTTCATG ACCTGTCCTC CCACATGATC CAGGCTCTAC CCTCAGGGCT
29551 TCATTACCTG TCTTCCCACA TGATCCAGGC CCATTCTTTC TTGAACCATT
29601 GGAAAGGAAT TTGCAGATAG GATGTGTACC CCTAACTGCC TGAGTATTTC
29651 TTAGCAGGTG TATTCTTTTG TGCAAGTGTA AGTCAGAATG TTAATGTTGA
29701 TGAAATACTA ATCTGCAGGC CTAATTGTTC CAATAATGTC CTTTATGGCA
29751 AAATCTTCCC CTCAAGCTTT AATCCAATAT CATGGGTTGC ATTTGTTTAT
29801 TTTTAATTAT TTTTCTTTTC TTTTTCTGTT TTCCTTACCC TTCTCACTGT
29851 GCACATGGGT TGCATTTAGT TATCACATTG ACACCCTTTT AACCTGGAAT
29901 AATTCCTTAA TCTTTCCTTG TGTTTGATGA CTTTGTCATT TTTGAATTGT
29951 TCCACAAGTT ATTTTGTAGA ATATCCTCAG TGTTTTTTTT TTTCTGGTGT
30001 CTCGTGATTA GATTCAGGTT ATGAAACTAC ATTTTTGTCA GGAAGATTGC
30051 AGAAGAAATG GGGCCTTCTC CTGCACCTTA CCAGGAAGCA CACACCAACT
30101 TTGATCCCTT GATTAAGGTG ATGACCGCTG TACTTAGTTT CTCCGCTATG
30151 AAGTTGCTTT TTTTTTCTTT GTGGGGAGAC AGTTTTAGAC TATGTAAACG
30201 TCCTATTTCT CATCATACTT ATACCTGTTA GTGTTAGCAT TTGATGATGA
30251 TTCTTGCCTG AATCAATTAT TTGCATAATG CTTACAAAAT TATCATTCCT
30301 TCCATATGTA TTAGTTGGTG TTCTTCCAAA AGCTTTTCCT TTGCCTCTGT
30351 TTATTTACTT ATTTATCACT GTGGACGCAT AGATTCTTAC GCAATTGATT
30401 TTGATACTGA TCTCATAGCT AGACAATTTT GCTAAACTTT TAAAAAAATT
30451 TATGTACTTT ATCTTTTATA GCAGCTTTAA ATTTACAGAA AATTTGAGTG
30501 GAAGATGCAG TGTTCCCATA AAGCCGCTAA CTCCTCGCAC CTTCCCTCAA
30551 GTTTCCCCAG TACTAACATC TTGCATTCAA GTGGTGCGTT TGCAACATTC
30601 ATAAATTATT ATCGTCCAGA GTCCATTGTT TACATTCAGC TTCCTCTTCA
30651 TGTTGTTCAT TCTGTGGTTT CACAGATGTG TGATGCATGT GCCCACCACT
30701 GCAGTGTCAC ACAGGATCTC ACTGCCCCGG AGTCCTCTGC GCTGTCCCCG
30751 CCTCCAGAAC CCCTTAGTAG CAAACACTGA TATTTTTACT GTCTCCATAG
30801 TTTTGCCTTT TCAGACTGAC CTATTCACT TAGTAAGAAG CATTTAAGAT
30851 TCCTGAGTCT CTTTCTATGG CTCAATAGCA CATTTCTTTT TAGTGCTGAA
30901 TAATATTCCA TTGTCTGGAT GTACCACAGT TTATTCATTC ACCTACTAAG
30951 GTGAATGTCT TGCTTGCTTC CAAGTTTTGG CAACTATGAA TAAAGTTGCT
31001 ATCAATGTTA GCGTGCACAT AAGTTTTCAG CTCATTGGG TAAATGCCAA
31051 GAAGCATGAT TGCGGGATCC TATGGTAAGA GTGTGTTTAG TTCTGTAAGA
31101 AGCTGCCAAA CTGTATCTTA AGTGGCTGCA CCATTTGCGT TTCCACCAGC
31151 AATGATGAGC GTTTTGTTGC TCCACATCCT CACCAGCATT TGCTGTTGTG
31201 TTTTGGGTTT TAGCCTTTCT AAGAGGTGTG TAGTGGTATC TCCTTGTTTC
31251 AATTTGCAAT TCCCTAATGA CATTATGTTA AAATCTTGTC ATATAGTTAT
31301 TTGCCATCTG TGTATCTTTT TCAGTGATGT GTCCTTTAAA GTCTTTGGCT
31351 CATTTTTAAA TTAAATTTTC TTATTGTTGA GTTTAGTTC TTCATATATT
31401 TTGGCTGCCA GTCCTTTATC AGATATGTCT TTCGCAAATA TTTTCTGCCT
31451 GTGTCTTGTC TTTTCATTCT ATTAACAGTA TCTTTTGCAG AGCCAGTTTT
```

FIGURE 3J

```
31501 CATTTCAAGG AAGTCCAGCT TATCAATGTT CTCTTTCATG TATCATGTTT
31551 TTGGTGTTGT ATCTAAAAAG TTACTGCCAA GCCCAAGGGT ACCTAGATTT
31601 TTTCCTGTGT TATATTCTAG GATTTTTAAA GTTTTGCATT TTACATCTAG
31651 GTCCATGATT CATTTTGAGT TAACTTTTGT GAAGGGTTTA TGGTTTGTGT
31701 CTAGATTTTT TTTTTTTTTT TTTTTTTGCA TGTGGATGTC CAGTTGTTTT
31751 GGTACCATCT GTCAAGAAGA CTCTTTTTGG GTCATTTTGT TGCCTTTGTT
31801 TCTTTGTAAA AAATCAGTTG ACTGCATTTG CATGGGTCTA TTTCTGAGCT
31851 CTCTGTTCCA TTGCATTGAT CTGTTTGTTC TTCTCAGCAA TCCCACACTG
31901 TCTTGGTTCC TGTAGCTCTG TAGTAGGCCT TGCAGTCAGT TACCGCCCCT
31951 GTTCTCACTT CAGTGTTCTC TTCAATAGTG TTTTGACTAT TCTAGGTTTT
32001 TTCCCTCTCC ATATACATTT TAGAGTCAGT TTGTCAATAG TTTACAAAAT
32051 AACTTGCTGA GACTTTGATT GGGATTACAT TGAATCTGTA GCTCAAGTTG
32101 GAAAGATCTT TTATTTCTTT CATCAGAATT TTGTAGTTTT CATCATATAT
32151 AGATCTTGTA CATATTTTGT TGTTTATACC TAAGGATTTC ATTTTTTTGG
32201 TGCTAATGTA AATGGCGTTG TGTTTTAAAT GTCAAAATCT AATTGTTCAT
32251 TGCTGGTAGG AAAACAACTG ACCCTTTTTT TTTTTTTTAA GGGACGCAGT
32301 CTTACTCTGT TGCCCAGGCA GAGTGCAGTG GTGCCATCAT AGCTCACTGC
32351 AGCCTCAAAC TCCTGGGCTT AAGGAATCCT CCTGTCTCAG CCTCCTGAGC
32401 AGCTAGGACC ACAGGCATGT GCCACTACGT TCAGCTAATT TTTCAATTTT
32451 TTTGTAGAGA TGGGATCTTG CTCTGTTGCC CAGGCTGGTC TCAAACTCCC
32501 GTCTGCTTTG AGATGATTAT ATATTTGTGT CCTTTGTTAA TTTAGAGGAT
32551 TATTATGGAT TTTTCTAATG TTAAGACACC TTTGTATTTC TGAGATCGAC
32601 CTTAGTATTG GTCTATATTT AAGACAGTAT TCAGTTCTC AGTTGTTTTT
32651 TGTTTTTTGG TTTTTTTTTT TGAGACAGAG TCTCTGTCTC CCAGGCTGGA
32701 GTCCAGTGGC ACAATCTCAG CTCACCGCAA GCTCTGCCTC CCGGATTCAC
32751 GCCATTCTCC TGCCTCAGCC TCCCGAGTAG CTGGGACTAC AGGCGCCTGT
32801 CATCATGCCC AGCTAATTTT TTGTATTTTT AGTAGAGACG GGGTTTCACC
32851 ATGTTAGCCA GGGTGGTCTC AATCTCCTGA CCTCGTGATC TGCCCACCTC
32901 GATCTCCCAA AGTGCTGGGA TTACAAGGCG TGAGCCACTG CGCCCGGCAG
32951 CAGTTTCTCA GTTTTAATTT GGAGTTTTGC ATCTGTGTTC ATGAGTGAGC
33001 CTGAAATTTT CACTTTTCCA TATCTTATTT CTCTGGGTTC CTAGAATGAG
33051 CTAGAGAGTG TTCCTCCTTT CTGTTCTCTG GAAGAGTTTG TGTGAGATTA
33101 GAATGAGTGT GTCTGATAAT TTAGTTGCAT TCATTTATAA AATTCCTAGG
33151 CCTAGAGTTT TTTTTCTGGG AAAAGTTTAC ATTTTGACTC ATTTTTTTAG
33201 TAGTTTTAGG ACTGTTTAGG TTCTCTATTT CTTGATTGAG CCAGTTTTGA
33251 TAAGTTAATC TTTCTAATTT GTAGATATTT TCTCTAAGTT TGCAAATGTA
33301 ATACATAAAA CTTTCTTGTC ATTTCTCACC ATATCTGTAG TTCTATCTTT
33351 TTATTGCTAA TATTACTAAT TTGTACTTTG ACTATTTGTA TTTGTTACCT
33401 GTTGCCGAGT AACAATATTA GTACAAACCT AGTGGCTTAG AACAACACAC
33451 ATTGATTACT TCACCGTTTC TGTGTGTCAG AAGTCCAGGC GCGGCCTCGC
33501 AGGTCGTCCT CTGCCTCAGG GTCTCTCCGG GCTTCAGTCA GGGTGTTAGC
33551 CAGGACCGGG GTCTCGCCTG AGCTTCCAGT GAGGAAGGAT CTGCCTCTGA
33601 GCACACAGGG TCCTCGGCAC GATCCCATTC CTCAGCTGGA AGCTGCCGAC
33651 TGCCGTCTGC TGCGGGGCCT CTCTAGATGG CATCTTCACA AAAGCGAGAA
33701 GGGAGAGTTG GTAGAGGGAG TCTGCTAGCA CCATGGGAGT CGCGGTCACA
33751 CAGACCTCGG TCCCAGGACC CGCACCCATC AACCCTGCCG TGATCTGCTG
33801 GTTAAAGACA AGTCCCACGT CCCACAGGGT GACACTGGAG TAGACACTTC
33851 GCTCTGGCCT TTTCAGAGAA CTGGTTATTT TTTGGAAATA TCAGTTAGAT
33901 GTAGGATGGG TCTTGTCTTC TAAATCTATT GTTTTCTCTC TAATTGATTT
33951 TTTCCTGTTT TTATTTAGTT CACTTTGTTG GGTTTGCTCA AGCCTGGGTC
34001 ACTGGATCTC AGGGATGCTG CTCCTGTTTG CAGCTGTGTC TGCAGGGGCT
34051 TCCCAAGGCC TTGCTTTCCC CTCACGTCCC TTTCTCAGAC TCTGCCAATC
34101 CGCTTCCCGC TCTGGTGTCC TGTGGTTGCT TCTTTTTAAA ACCCTCATCG
34151 GTCTGTGTAA ACTGTTTATT TTTATGTGG TTTTAAGGGA GACCATTCTC
34201 ATTCTTTTGA GACCCTGGAA AGGATGGAAT TGGGATAGGT AAACTGCTGT
34251 TTTACCAGAA TGTTCACTGG ACCAATCTCG TGTTCCAGGG AGACCCTCAC
34301 GCAGGGCTTA GAGTTCTGTC GGCGGCAGCA GAGGGCCGAT GGCTCCTGGG
34351 AAGGGTGAGT GAGCCTCCAC TCGTGAGTGC AGAGATGCAT GGGATCCAGA
34401 GGTTTCTGCT CTCACACACT GCGTTCATAA ATGTTGGCTT GTATGTTGTT
34451 GCTACACCAG AAGTTTCTGG AAGTGAGCTG CCAGCCCGTG ACTTCTGGGG
34501 GACCTCGTTC CTTTGTGGCA TGCGTGGCCT TTGCCCCGGT GGAAATTGCT
34551 CAGTACGTTG CTGGGCGCAG CCGGGCTGCT GGGAGCGCGC TGTAGCCTGA
34601 GCGTGGCTAT TCCCTCCACC CTTTCTGCTT GCTCTTAGGG TCCAGCAGAC
```

FIGURE 3K

```
34651 AGAGCTGCTG TCTTCCACGG CCTTAATGCC TGAGGCACTG GAGTTGGTGG
34701 GCTGGCTGGG GCACGTGTGA TTGTTGCAGA ATGCGTGTTG TTTCACACAC
34751 CGGCTGTGAA CAGGGTGGAA GGGCTGAGGC TCTCCCTGTT TCCCTCCAGC
34801 TCCTGGGGAG TTTGCTTCAC CTACGGCACC TGGTTTGGCC TGGAGGCCTT
34851 CGCCTGTATG GGGCAGACCT ACCGAGATGG GTGAGTGAGT GCCTGTCCTC
34901 TGGTGGGTGG GGGTTCTCAA CCCAATGCTC TGTCATGAGT GTTTTTTGCT
34951 TTGACATTTG GTTTTAGGGT TTGTTTGTTT GTTTGTTTGT TTTTGAGACG
35001 GAGTCTCGCT CTGTCAACCG GGCTGACATG CAGTGGCATG ATCCTAGCTC
35051 ACTGCAGTCT CAAACTCGTG GGCTCAAGCG ATCCTCCCGA GTAGCTGGGA
35101 TCACAGGTGC ACGCCACCAC CCCGGGCTAA TCTTTTAAAA CTTTTTATGTA
35151 GAGATGGAGT CTTGCTGTGT TGCTCACACT GGTTTGGGCT CAAGCAGTCT
35201 TCCTACCTCG GCCTTCCAAA GTGCTGGGGT TACAGGCATG AGCCAATGTG
35251 CCTGGCCTGT TTTTAATATT TTTAAACAGT GAGATAAGAT CCCCGGTTGA
35301 AATGAAGATG TTTCCCTGGT CCCACAGCTC TCTGGAGCTT CCTGACATGT
35351 ATGCTGGAGG GACGCTTCTG GTCTCCGGCC CCTCCAGGCA TACAGATGCC
35401 TCCCAACCCT GAGTAGGAAG ATTAGGGTCC ACGGCCTCGC TGGAGCGGGT
35451 TAGAAGGCAG GAGATCTCCG GTCCCAGCCG TGTCTCCAGC CGCCGGACTC
35501 TCTCCCAGCC CTGTCTCCAG CTGCCCCACT GTCTCCCAGA GTCTGCCGTG
35551 TGGATGTTTA GAGGTGGGGA GCACCGTGCT TGGCTGAGTG CAGCTTGTGA
35601 GACGCTGCTC CCAAGCACTG CAGACCTCAC TCAGCCTGAC GCGTCCGTGA
35651 GGCCATCCTC GGTACTCGCA TGTCCCTTTG TCTTCCCAGC GACTCTGGGA
35701 GGCAGGAGTA TCTGTTCCCA GTTCACATCT GCAAAAGTCA AGCTCGGGTT
35751 TCAGTAGTGG CCCATGGCCC TTAGGTAGGG TGGCCCCATC GTGCAGGCTC
35801 CTCCCCGTAC CCCAAGGCAG CCTGCTGGGG TGAGAAGCCA GGGGTCTGGG
35851 ACCTTCCTTG GTGTGATGGT GTCTCCTGTC TCTGGTCTTT GCAGGACTGC
35901 CTGTGCAGAG GTCTCCCGGG CCTGTGACTT CCTGCTGTCC CGGCAGATGG
35951 CAGACGGAGG CTGGGGGGAG GACTTTGAGT CCTGCGAGGA GCGGCGTTAT
36001 TTGCAGAGTG CCCAGTCCCA GATCCATAAC ACATGCTGGG CCATGATGGG
36051 GCTGATGGCC GTTCGGTGGG GACGACGGGA CCGTCCCTGA GCCTTGGGTT
36101 TGGGTAGAGG AGGGACACTC AGCTGTGAGC CGGTGGCCTG GGCTGAGTGA
36151 ATGTAGAGAG GAGGGGAGGC CTGTGGGCCA GGTCAGCTGC CACTCTGGGA
36201 ACAGACACCT ACAAGAGCCA CATGCCTGGT TCCTGGGGCA AGAACGTGGG
36251 CTGCTCTGAC CAAGTGGGGC CCTGCAGAGA GGCTCGCCTC TTAGAAGTGA
36301 ACCACCCACC ATTAGCCATG TCAGTGGAAG AGCAAGCACA TCAGGGACCC
36351 ATGGAAACAG CGAGGTGGGC TGCGATGAGG ATGCTGCTTC CTGGTGTGGT
36401 AGTGATGACG GTCACAGCAG CTGCTCTCTG TGGCCCTACT GTGTTCACAG
36451 CTGGTGCTGA GCCACATATG TGCCAGGTGC ACACACACGC AGACGCATGC
36501 AGGCAGGCAT CAGTGTACAC ACTGATGTGC ACACACAGAT GTACATGGAG
36551 ACAGATGCAC ACACAGGCCT ATGCACACAC GTACGCATGC CCACACAGGC
36601 ACCTGTGTCC ACACACATAC AGATGCACAC ACAGCATCCC ATCTGTGCCA
36651 CACACTGACA TAGGTACATG GAGACAGATG GAGACAGATG TCTGTGCACA
36701 CACGTATGCA TGCACAGGCA CCTGTGTACA CACACGTACA GATGCACCCA
36751 CAGGATCCCA TCTGTGCCAC ACACAGACGT AGGTACATGG AGACAGATGC
36801 ACACACAGGT CTGTGCACAC ACATACATAC GCATGCACAG GCACCTGTGT
36851 ACACACATGC AGATACACCC ACAGCATCCC ATCTGTGCCA CACACAGACA
36901 TAGGTACATG GAGACAGATG CACACACAGG TCTATGCACA CACATACGCA
36951 TGCACAGGCA CCTGTGTACA CACACGTACA GATGCACCCA CAGGATCCCA
37001 TCTGTGCCAC ACACAGACGT AGGTACATGG AGACAGATGC ACACACAGGT
37051 CTGTGCACAC ACATACATAC GCATGCACAG GCACCTGTGT ACACACACGC
37101 AGATACACCC ACAGCATACC ATCTGTGACA CACACAGACG TAGGTACATG
37151 GAGACAGATG CACACACATG TCTGTGCACA CACATACATA CGCATGCACA
37201 GGCACGTGTG TACACACATG CAGATACACC CACAGCATGC CATCTGTGAC
37251 ACACACAGAC GTAGGTACAT GGAGACACAT GCACACACAG GTCTGTGCAC
37301 ACACATACGC ATGCACAGGC ACCTATGTAC ACACATGCAG ATACACCCAC
37351 AGCATCCCAT CTGTGCCACA CACAGACATA GGTACATGAA GACAGATGCA
37401 CACACAGGTC TATGCACACA CGTATGCATG CACAGGCACC TGTGTACACA
37451 CATGCAGATG CACCCACAGT ATCCCATCTG TGCCACACAC AGACATACGT
37501 ACATGGAGAC AGATGCACAT ACAGGTCTAT GCACACATGT ACACATGCAC
37551 AGGCACCTGT GTACACACAT GCAGATGCAC CCGCAGTATC CCATCTGTGC
37601 CATACACAGA CATACGTACA TGGAGACAGA TGCACATACA GGTCTATGCA
37651 CACATGTACA CATGCACAGG CACCTGTGCA CACATATGCA GATGCACCCG
37701 CAGTATCCCA TCTGTGCCAC ACACAGACAT ACGTACATGG AGACAGATGT
37751 ACACACAGGT CTATGCACAC ATGTACACAT GCACAGGCAC CTGTGTACAC
```

FIGURE 3L

```
37801 ACATGCAGAT GCACCCGCAG TATCCCATCT GTGCCACACA CAGACATACG
37851 TACATGGAGA CAGATGCACA CACAGGTCTA TGCACACATG TACACATGCA
37901 CAGGCACCTG TGCACACATA TGCAGATGCA CCCGCAGTAT CGCATCTGTG
37951 CCACACAGAC ATACGTACAT GGAGACAGAT GTACATACAG GTCTATGCAC
38001 ACATGTACAC ATGCACAGGC ACCTGTGCAC ACATACATAC AGATGCACCC
38051 GCAACATCCC GTCTGTGCTG CCCTATTAGG TTTGTGGCCA TTTGGGGAAT
38101 CTTCCTAAAA CCCTAAAAGC TAGGGCAGGT CTGCTTGAGC AGGAGCAGCA
38151 GGGTCTGGGG GACCCCTGAG GGCAGGACAG TCAGGGACCC ACAGTTGAGC
38201 TGGGCCCGCT GAGCCCTGGA TCCTTCTTGG TGTCTTATCC TGGCCAGCAA
38251 GCAAGTGTGA GCTCCTGTGG GTCTCCAGAG GCCCATGAGG ACCAGTGGGC
38301 CAGTTGGGAA CAAGGCTTGG CGTCCTCTTC AGGGGGGAAC ACCAGGGCAG
38351 GCCTGAGGAG GCCTGTGTCC CCAGCCTGTC ATTGCTGTGG CTCCGCTTCT
38401 CAGGGAGCCT AGGAAGAAGG TGTGGCAAGA GCCCGAGGCG CTGGCTGCAC
38451 CTGGCGGGGC CTGTGGGCGT CAGTTTAGAC CCATCCATTC TCACTGCAGC
38501 ATTCCAGGGT TTGCCCTTAT GCTCGGCTGT GTGAGGGTGA GGATGATGCT
38551 GTGGGGGCAT GCATGCTGGG TGTGTTTCAG CCTTCTCTTC CACCAGGCAT
38601 CCTGACATCG AGGCCCAGGA GAGAGGAGTC CGGTGTCTAC TTGAGAAACA
38651 GCTCCCCAAT GGCGACTGGC CGCAGGTATG CCGCCAGGGA CCTGAGCGCA
38701 CAAGGCCCAG CACTGACCTC CAGCGTGCAT GGCTGTTTCC ACGTCCCCCT
38751 GCTCTGTGTC CTTTTTGGGG TACTTTGGAC ACTTGGGAGG CGTCACCTCT
38801 GCCAGTGAAT GCCACAGTTG GTGGCAGGTC TGTGGCAGGT GGTCGGGTCC
38851 TAAAGTCCAG ATCTTGCTGT TGTTTCAAGT GATGCTCTGG GTGGGGGAGG
38901 AGCTGGATGG GAGAAGCCAG TGGGCGGGAA GCCTTTTTGC TGCAGGACAG
38951 ACCCTCCCAC TCCAGATGAC CTAGTGGCCC CTCACTGAGC CAGAAGTCCC
39001 TGTGGTGTGG GTGTCATGAG GTCATGTGAG GCCAACCGCC CTCCCCTGGG
39051 ATGAGGCTGA GTTGGTGGAA GCTGATGTGG TTGTGAGGGG CTGGTGACCC
39101 TGGCTTAGGG TTTGCTGCAG GGCGGGGAGT CTGAGCTGGG CTGATGGTGC
39151 CATGACTGAT GCGGGATGGA CTACTTGCTT TCCTATGCTC TTGCTTAATT
39201 AGCCCTTTCC AGGCTGACTC ACCCACAAGC CAGCCAAGCC AACAGCCAGG
39251 GCTCCAGTTC AGGGACTAGC CCTCAGCTGA CTGGTGAAGC CTTTGTGTTT
39301 ATTTCTCTGT GTTCTTTTAG GAAAACATTG CTGGGGTCTT CAACAAGTCC
39351 TGTGCCATCT CCTACACGAG CTACAGGAAC ATCTTCCCCA TCTGGGCCCT
39401 CGGCCGCTTC TCCCAGCTGT ACCCTGAGAG AGCCCTTGCT GGCCACCCCT
39451 GAGAACATGC CTACCTGCTG GGTGCCGTCT GTGCGTTCCA GTGAGGCCAA
39501 GGGGTCCTGG CCGGGTTGGG GAGCCCTCCC ATAACCCTGT CTTGGGCTCC
39551 AACCCCTCAA CCTCTATCTC ATAGATGTGA ATCTGGGGGC CAGGCTGGAG
39601 GCAGGGATGG GGACAGGGTG GGTGGCTTAG ACTCTTGATT TTTACTGTAG
39651 GTTCATTTCT GAAAGTAGCT TGTCGGGCTT GGGTGAGGAA GGGGGCACAG
39701 GAGCCGTGAC CCCTGAGGAG GCACAGCGCC TTCTGCCACC TCTGGGCACG
39751 GCCTCAAGGT AGTGAGGCTA GGAGGTTTTT TCTGACCAAT AGCTGAGTTC
39801 TTGGGAGAGG AGCAGCTGTG CCTGTGTGAT TCCTTAGTGT CGAGTGGGCT
39851 CTGGGCTGGG GTCGGCCCTG GGCAGGCTTC TCCTGCACCT TTTGTCTGCT
39901 GGGCTGAGGG ACACGAGGGC AACCCTGTGA CAATGGCAGG TAGTGTGCAT
39951 CCGTGAATAG CCCAGTGCGG GGGTTGCTCA TGGAGCATCC TGAGGCCGTG
40001 CAGCAGGGAG CCCCATGCCC CTGGGTCGTG AGCTTGCCTG CGTATGGGGT
40051 GGTGTCATGG AGCCTCATGC CCCTGGGTCG TGAGCTCGCC TGAGTATGGG
40101 GTGGTGTCAT GGAGCCGCAT ACCCCTGGGT TGTGAGCTCG CCTGCATATG
40151 CAGGGTCTGT CATGGAACAT CCCAAGTCTG TGCAGCAGGG GAGCCCCATG
40201 CCCCTGGGAC ATGAACCCAC CTGCGTGGAA TGCTGTTTGT GAGGTGTCTA
40251 CAGGGTTTAT AGTAGTCTTG TGGACACAGA AATGCACAGG GGACACTTAC
40301 GGACACAGAA ATGCACAGGG GAGGCCGAGC ATAACCAGGG GTGAGGGGCA
40351 GGCAGCAGTT GTAGTTACTG CCGCGGGGCA CTGCTATGTG CAGGGACAGC
40401 CAGCGCCCAG CCCATCACCA CTCCCTGGGC TGGCTGGCAG GTATGGCACC
40451 CTGGGAGCCC GGCATATACC CAGGGCACCC CTACGGCTGC CGCCAGTCTC
40501 ATGCCCAGGT GGGTGCTCTG GGCTGGAGCG AGGGCCAGGT TTTGGGCCGA
40551 GGCTTCCCCA GGCAATCCTG TGAGCTCCCT TCTAGCCTCT GACCCAGTCT
40601 GGTCTGGCTT GCATGGATGT AGGGCTTGGG GTGGGAAGTT CAGGTCCTGG
40651 CTTTGCCTTT GCCTGATGTG GATGAGCAGC TCACATGCTC AGGGCCACCT
40701 GAGACTGTCA CTGCTCTCCC CTGGCTACTG GGAGGAGTCA CTGAGAGCTT
40751 CGTTACCCCT GCTGCCTTGC CCAGGGCACA CCCTATACCT CCTCATCTGC
40801 TCTTCCCCTC CCTGCCGCCT TCTGGGCAGG TAGCAGTCCC TGGCCTCTCC
40851 CCCTGGCTGA TCACTCTCCC TCAGGCAGTG GAGATCTGCG TCTGGACACC
40901 CTCAGATCCT GTCATTGCCT GCCCAGAGTC CTTCAGGGGC ACCCCTCTGC
```

FIGURE 3M

```
40951 CTTGGTGTGC GGTCCAGGGC TCTCACCCAG GTGCCGCACC CTCTGGGGTC
41001 TTCTGTCCAG CTCCCTTGCC CCATGTGCTG TCACTGACTC TCCTTGGGAC
41051 TCGCCTGCCT GCTCAGAGCC CTGCAGGGCT TGGTCAGCTG CCTGTTCAGT
41101 GTCAACACTT CCCTGCACAT CTTAAAACTG GGCTTTATTT TCGCTGAAGG
41151 AACTGTGTTG GGACCCTTGA CATCTGTCAG GTTTGCACAT GCTGTTTTTT
41201 TTTCTCAGCC CACGTGTTCT CCCCCACGTG GGGTAGCAGC AGGACAGACA
41251 GTGAATCACA GAGTCTGCCC TGAGCAGAGG CTGCTGTCCC TGGGACTCCT
41301 AGCCATGGTC AGACTGTACA AAACGGTTTT CCAGAAATGA AATGTAAATC
41351 CATTTTTATA CTGAAAATGT TACTGAAAGT CACTTTTATG AGCATCTGCC
41401 TTAATAAACA GACATTGATT CCCTTATCAG AAGCCTGTCA CACTGTGTTT
41451 CGTTTCATCC TGGGGAGAAC TGCAGATTTG GGGTTTCTGG CTGTCATACG
41501 TCACCTGCCT GTGGGGCGAG TGGGAGGCCC AGCCTGGTTT AGGGAACAAG
41551 AGTGACGTGA GGAGTAGCAG GGTGCGTCTC CAGTTACCTG AGGGAAAACA
41601 GATATTTTAA GAGATAATAG CATAGCCTAT TTTAATATGT TTTAAAGGCC
41651 ATAAGCATAT CCAGGAAGAT AAATAAACGT GATACAATGT CCACATAGGA
41701 GGAACTTTCT TTCACTGCAT TGTTTTCCTT CACAGTGGCC TTCAAGTCAC
41751 AGGACGCAGC GATTCCCTGC CCTCTTCGGT GTTATTACAC AGGCAGGACT
41801 TCAGTGTCAG TATCCCTGCC TTCAGTCTTC TTTAGAAATC ACATCTGTGT
41851 TCAATCCATT GTTTAGAGGG AGTGTATTTT TCCTGTTCCA CGAAGAGGAC
41901 TTTTTGTTCA CAATTGGATC ACAATGCAGA GGAGTCTGTT CCTCCCCCGT
41951 CGGCTTCTCG GTGCTGGGAG GGTGACCTGT CCCAGATGAC TCATCACCCT
42001 GACATGCTCT TGACAAAGGA CACCACCAAG AGGAGATGGC AGCTGTACCG
42051 GTGCAGCCTC TGTCTGAGGG GGATATTTGC CTCAGTGTGA TTAAAAATCA
42101 GTCATGAAAG ATTTTTGAAT TCAGATTATT TTTATCAGGA ACAGATTTTG
42151 AACATCCTGA AATCTTTTCC CTGGCATCAT ATTAGGTTTT CTTTGTTCAC
42201 TATGATGTAA AGTTTCAGAC TCTTGATATT TTTAATATCA ACATAGACGG
42251 TAGGACAAGG AACGGTACCA GAAATGAGTA AAGAGACAAT AATGATAAGA
42301 TCGATTTATC AAGACATAAC AACCCCAAAT GTATATGCAC TAAATAACAG
42351 CTTCAAAATA CATGAAGCAA AATGGCAGAA TTGAAGAGAA TGAGATAAAA
42401 ACAGAATTTT AACGGGTGCT TTCCGTACTT TGTAACTGAC AGACATGAGA   (SEQ ID NO:3)
```

FEATURES:

```
Start:    2034
Exon:     2034-2047
Intron:   2048-2179
Exon:     2180-2345
Intron:   2346-3088
Exon:     3089-3227
Intron:   3228-8043
Exon:     8044-8119
Intron:   8120-8806
Exon:     8807-8928
Intron:   8929-11095
Exon:     11096-11192
Intron:   11193-14163
Exon:     14164-14299
Intron:   14300-14894
Exon:     14895-15003
Intron:   15004-15390
Exon:     15391-15509
Intron:   15510-16853
Exon:     16854-16951
Intron:   16952-17636
Exon:     17637-17664
Intron:   17665-19945
Exon:     19946-20002
Intron:   20003-21064
Exon:     21065-21136
Intron:   21137-22389
Exon:     22390-22440
Intron:   22441-23113
Exon:     23114-23263
Intron:   23264-23922
```

FIGURE 3N

```
Exon:      23923-24019
Intron:    24020-24749
Exon:      24750-24855
Intron:    24856-34288
Exon:      34289-34354
Intron:    34355-34799
Exon:      34800-34880
Intron:    34881-35894
Exon:      35895-36065
Intron:    36066-38596
Exon:      38597-38675
Intron:    38676-39320
Exon:      39321-39449
Stop:      39450
```

CHROMOSOME MAP POSITION:
Chromosome # 21

ALLELIC VARIANTS (SNPs):

DNA Position	Major	Minor	Domain	Protein Position	Major	Minor
478	–	A	Beyond ORF(5')			
891	C	G	Beyond ORF(5')			
948	–	C	Beyond ORF(5')			
3311	A	T	Intron			
3616	T	C	Intron			
3910	G	A	Intron			
6028	G	A	Intron			
8299	G	A	Intron			
8373	C	G	Intron			
8424	A	G	Intron			
8680	A	G	Intron			
8700	C	G	Intron			
8996	A	C	Intron			
10590	T	C	Intron			
11090	G	C	Intron			
11710	G	A	Intron			
12591	G	A	Intron			
13431	–	T	Intron			
14746	C	G	Intron			
14975	G	C	Exon	277	P	P
16031	C	T	Intron			
16891	–	T	Exon	339		V
19359	C	T	Intron			
19405	A	G	Intron			
19653	G	A	Intron			
19742	T	C	Intron			
20054	A	G	Intron			
20627	–	A G	Intron			
21337	T	C	Intron			
21894	C	T	Intron			
23360	G	T	Intron			
26758	A	C	Intron			
27033	T	C	Intron			
27332	C	A	Intron			
27538	C	A	Intron			
27625	G	C	Intron			
27736	A	G	Intron			
30688	T	C	Intron			
31172	C	T	Intron			
31433	C	T	Intron			

FIGURE 3O

32660	G	T	Intron			
32981	A	C	Intron			
33557	T	C	Intron			
33652	G	A	Intron			
34390	T	C	Intron			
34399	G	C	Intron			
34989	G	–	Intron			
35067	C	G	Intron			
35495	G	A	Intron			
36001	T	G	Exon	631	L	V
38948	C	T	Intron			
39160	T	C	Intron			
40405	G	A	Beyond ORF(3')			
40794	C	T	Beyond ORF(3')			
40961	A	G	Beyond ORF(3')			
41891	C	T	Beyond ORF(3')			

Context:

DNA
Position

478   AGGTTCAGTGTGAGATTCCATCCAGGCTGAAGCCCCTTATCCCTATTCTTCATGTTTCTA
CATGGAGGAACTTACCTGGAGAAAAACTTCCAGCCTCTTTCTGCTTCCAGAGAAGTAGAG
TGACTCATTTGATTGAATTTCAGAGAACAGATAGGGTGGAGTGTGCTCAGGCTCCTCTGG
GTACTCTTTCTGGGGTCTGTGGGTTGACTGGAGGGGTGTCTTCTGGTGGGCACTCAATTG
CATAGTGCTTGGTGAGGCAGTTTCATGGCCTAGAGGCTGGGGGATATGTTTGTCTGACTT
[-,A]
CGGGTGATTTAGTAGCTTGCCCTCTTGCTTGCAGATTTAAGCCTTGTCCTTCAAGCTAGG
TTTTTAATTTGTGGCAAAGCTGATATTTTGATACCCACCCATCTTATTGCTGTGTCTTTT
TCATCCGTTTCTGAACTGGGATAGGAAGAGGTGATTATCCTTGATTGTCTAAAACCCCGC
TATTCCACTGTGGGGAAGGTGCCTGTGGGTATTCTTTTGTCCACTCTCTCTTCCAACTTT
CTCCTCCGGCTTGCTGTGGCTCACCGCCCCTTCGAAGTTAGGCTGGGGGTAGGAATTGAG

891   TGTCTTTTTCATCCGTTTCTGAACTGGGATAGGAAGAGGTGATTATCCTTGATTGTCTAA
AACCCCGCTATTCCACTGTGGGGAAGGTGCCTGTGGGTATTCTTTTGTCCACTCTCTCTT
CCAACTTTCTCCTCCGGCTTGCTGTGGCTCACCGCCCCTTCGAAGTTAGGCTGGGGGTAG
GAATTGAGGAGTGGGTGCCGAAATGCTCACTAGGCTGGGGCAGTTGTAACTGGATGTCAG
GGCTTCTGTGGGCCAGGTGAAGACATGCTGGGGTCTTCTGTGGGTCCTTGACCTGACTTA
[C,G]
GGACCACTGGCTGCAGCCTCCAGACGTCAGCCATGTTTCCAACAGTCAGACGCCCCCTGC
CCTGTTGCGCCCGGCTGTCCCTTCCAAGTTCGGTCACTCGCTCTGCCTCCATCTTCCTCT
TCCCTCTGCTGCTAAGGCTTTTCACCTTTAATTTCTCCTGGGGCCACCCCCAACTCCAGC
GACCCCGTGAGCAGCTGAGGCTCTACCGCGCTCGGTCCTGGCCAGCGACGCAGCCCTTCC
CTGGCGGGGCTCCAGGGCTTCTGGCCCCTGTGGTCCGCCAGGTGTGGGGGCCCACGGCCT

948   TAAAACCCCGCTATTCCACTGTGGGGAAGGTGCCTGTGGGTATTCTTTTGTCCACTCTCT
CTTCCAACTTTCTCCTCCGGCTTGCTGTGGCTCACCGCCCCTTCGAAGTTAGGCTGGGGG
TAGGAATTGAGGAGTGGGTGCCGAAATGCTCACTAGGCTGGGGCAGTTGTAACTGGATGT
CAGGGCTTCTGTGGGCCAGGTGAAGACATGCTGGGGTCTTCTGTGGGTCCTTGACCTGAC
TTAGGGACCACTGGCTGCAGCCTCCAGACGTCAGCCATGTTTCCAACAGTCAGACGCCCC
[-,C]
TGCCCTGTTGCGCCCGGCTGTCCCTTCCAAGTTCGGTCACTCGCTCTGCCTCCATCTTCC
TCTTCCCTCTGCTGCTAAGGCTTTTCACCTTTAATTTCTCCTGGGGCCACCCCCAACTCC
AGCGACCCCGTGAGCAGCTGAGGCTCTACCGCGCTCGGTCCTGGCCAGCGACGCAGCCCT
TCCCTGGCGGGGCTCCAGGGCTTCTGGCCCCTGTGGTCCGCCAGGTGTGGGGGCCCACGG
CCTCACCGCGCCTACCCCACTCCCCCCGGCGAAGCTACGCGGCGCTCAGCTTCCCAGGGA

3311   TGTATGTGAAGAGGGTTCCTCTGGCCGGGCAACAGTCCCGTCAGCTATCTCTTTTTTTTT
TTTTCGATCTCTTTGCAGAAGAATTACTTTAAGGACTTGCCCAAAGCCCACACCGCCTTT
GAGGGGGCTCTGAACGGGATGACATTTTACGTGGGGCTGCAGGCTGAGGATGGGCACTGG
ACGGGTGATTATGGTGGCCCACTTTTCCTCCTGCCAGGTAGGAGTATGCTGCCCCAGCCT
GATGGTATGGCCACCCTGGATCACCCTTGGGATCCTGGCCCAGCCTGGTCTAGGGTTTTG
[A,T]

FIGURE 3P

```
      TGAAGCAGGTGAAAATCCAGGGGCTCACAAGAAAAGGGCTGGCAAACTCTGCCCTATGTC
      AGAGTCGTCCTGCTATTGGTCTAGGGGATCAGCTAGCCTTGCCAGTGTAGGGTGACAGGC
      TCTCTGATAAGAGAAGCAAGTGGTTCTCTAGGGCTCTGTGTTGCCTTGAGGGAGGAGGAA
      GGTGGGCTTTGAAGTCTCAGTACAGGATGGGATGGACATTCCAGGTGGAAGGCCCAGCCT
      ATGCCAAGGGGCTGTAGGTGGGCAGAGTGGTGGGTGGGGAGCTGATATCTGCTGTGAACT

3616  GCAGGTGAAAATCCAGGGGCTCACAAGAAAAGGGCTGGCAAACTCTGCCCTATGTCAGAG
      TCGTCCTGCTATTGGTCTAGGGGATCAGCTAGCCTTGCCAGTGTAGGGTGACAGGCTCTC
      TGATAAGAGAAGCAAGTGGTTCTCTAGGGCTCTGTGTTGCCTTGAGGGAGGAGGAAGGTG
      GGCTTTGAAGTCTCAGTACAGGATGGGATGGACATTCCAGGTGGAAGGCCCAGCCTATGC
      CAAGGGGCTGTAGGTGGGCAGAGTGGTGGGTGGGGAGCTGATATCTGCTGTGAACTTCCT
      [T,C]
      GGGGCTATTGCAGGAGAGCTTCAGGTTCAGGCTGGTGAGTAGGAGGAGCATAGCAGTTGG
      ACTGCCTGGGTATTGAACTGATTTGGCTACACAAGACTATTTTGCATCCTGGGAGTGTTT
      CTCTACAGAAATCCTCAGCCTTGTAAAATGGGAAATTCCCTCCTATGAATTTATGCAATA
      GGACTTTTTTCCCTAGTGACTTGTAATCACATTGTTTCAATGACGTGAATTCCTACATAA
      ATAGGTTTTGTTTCTGTGATAACTCTTACTGATACATCATTTTCTTTTACTACGCTGACT

3910  CTTCCTCGGGGCTATTGCAGGAGAGCTTCAGGTTCAGGCTGGTGAGTAGGAGGAGCATAG
      CAGTTGGACTGCCTGGGTATTGAACTGATTTGGCTACACAAGACTATTTTGCATCCTGGG
      AGTGTTTCTCTACAGAAATCCTCAGCCTTGTAAAATGGGAAATTCCCTCCTATGAATTTA
      TGCAATAGGACTTTTTTCCCTAGTGACTTGTAATCACATTGTTTCAATGACGTGAATTCC
      TACATAAATAGGTTTTGTTTCTGTGATAACTCTTACTGATACATCATTTTCTTTTACTAC
      [G,A]
      CTGACTTTGTAATAGATAGAAAGTCCTTATATACCTTTGTTGCCTTTCTTTTTAAAACAT
      CTCTTACCTGTGTCTATTCATTTACTCATCCAAATTGCCTTTATCCTGATTTTGTCCCAG
      ACTTGAAATGAAGTTGCAATAGGCTTATATGTTAGTTTGGGAAGAGTTGGCCTTTAACGT
      TAAAAACAGTTCCATGGTGTTTACTGTAGGCCAAGCCCTGCTCAAGGCCTGTTCTTCTTT
      TAGTCCTTAGAATAAGCCTAATGAGATACATTAGAAAGCTGAGGCACATTTATTCCAGGT

6028  GTCTCTTGCCTTGGCACCAAGGTGGCTTGCCACCCACAGCCTCTCGAGTAGCTGGGATTA
      CAGCCATGTGCCACCATGCCTGGCTAATTTTTGTATTTTTGGTAGAGACAGGTTTTCACC
      TTGTTGGTCAGGCTGGTCTCGAACTCCTGACCTCGTGATCCCCCACCCCCCACCCCCAGC
      CTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACTGCACCTGGCTGAGTTGGAGCTTTTC
      TTCCCTCTTTTTGGACTTTGGAAAATGCTCTTGGTCCATGATGCTATGTAGACAGCTCCC
      [G,A]
      TTGACTGTGGCCTGTGCGGCATTGGGCAGCACTCTGGTGAACACTGAATCGGGTCTGACC
      TCCTAGCCCCACCATTTACTGGCTGAGCCTCAGTTTCCTTGCCTGTAAAATCAGGAAGAT
      GCTGGCTCTGCTCCTCTCTGCACATTTCCCCGTCCTAACAACATTATAACTGTTAGGAAA
      GAGACGGGCTTGTTTTGGGATGGCTCATTTTATGTGACCCTGTGCGCTGTCTCTGAGTCC
      ATCTGCCCTTCTTCCAGGGTGTAGGGACCAGCCCCACAGGGTCGGTGGGTCTCTCCCTGT

8299  CATGACCACCTGTCCCCAGTGAGGAACATCTCTCCTGCCACACAGGCCTCCTGATCACTT
      GCCACGTGGCACGCATCCCTCTGCCAGCCGGATACAGAGAAGAGATTGTGCGGTACCTGC
      GGTCAGTGCAGCTCCCTGACGGTGGCTGGGGCCTGTGAGTGTGCCTGCCCCTGTGTCACT
      GCACATGTGCATGTGTGTGTTCTCATGATGTAGGAGATGCTTGGGTTTCCAGGCAGCTGC
      CAGGGGTTAGGAGTGATTGCAGCTGTGGGTGTGGGGTGGGTGAGGGAGAGACTAGCAGGC
      [G,A]
      GGGAGTGGGCTGAAGGCCATGCAGGTGGGGCCTCGGCTTCACATCTTTTGTTAAATGGAT
      TTTGTGGCTGTTACGACACTCTTGAGACCCACATGTGAAAACTGTCAGTCTGTTATCACT
      TAAGACAGAAGAAAATTGCCCTTGACTCTGGGCTGGCAGCAGGTGGAGACAAGGCCTGAC
      AGCTTTCCTGCCATGTGGCACACACTTTGGGAGCAGAGCCATAGCCCAAAGTGGACCGCC
      CTTGAGCTAGAAGTGTTGACTCAGGCGTGGGAAGGTGTAGAGCAGGCGGGTCACGGTGAG

8373  ATCCCTCTGCCAGCCGGATACAGAGAAGAGATTGTGCGGTACCTGCGGTCAGTGCAGCTC
      CCTGACGGTGGCTGGGGCCTGTGAGTGTGCCTGCCCCTGTGTCACTGCACATGTGCATGT
      GTGTGTTCTCATGATGTAGGAGATGCTTGGGTTTCCAGGCAGCTGCCAGGGGTTAGGAGT
      GATTGCAGCTGTGGGTGTGGGGTGGGTGAGGGAGAGACTAGCAGGCGGGGAGTGGGCTGA
      AGGCCATGCAGGTGGGGCCTCGGCTTCACATCTTTTGTTAAATGGATTTTGTGGCTGTTA
      [C,G]
      GACACTCTTGAGACCCACATGTGAAAACTGTCAGTCTGTTATCACTTAAGACAGAAGAAA
      ATTGCCCTTGACTCTGGGCTGGCAGCAGGTGGAGACAAGGCCTGACAGCTTTCCTGCCAT
      GTGGCACACACTTTGGGAGCAGAGCCATAGCCCAAAGTGGACCGCCCTTGAGCTAGAAGT
```

FIGURE 3Q

```
          GTTGACTCAGGCGTGGGAAGGTGTAGAGCAGGCGGGTCACGGTGAGGAAGGAGTGGGGGG
          CTCAGTTGTCATGGGAGGTGCATGAATTCGTACTGCAGAGTGGCTGCTCAGGGGTCTCCT

8424      GTGCAGCTCCCTGACGGTGGCTGGGGCCTGTGAGTGTGCCTGCCCCTGTGTCACTGCACA
          TGTGCATGTGTGTGTTCTCATGATGTAGGAGATGCTTGGGTTTCCAGGCAGCTGCCAGGG
          GTTAGGAGTGATTGCAGCTGTGGGTGTGGGGTGGGTGAGGGAGAGACTAGCAGGCGGGGA
          GTGGGCTGAAGGCCATGCAGGTGGGGCCTCGGCTTCACATCTTTTGTTAAATGGATTTTG
          TGGCTGTTACGACACTCTTGAGACCCACATGTGAAAACTGTCAGTCTGTTATCACTTAAG
          [A,G]
          CAGAAGAAAATTGCCCTTGACTCTGGGCTGGCAGCAGGTGGAGACAAGGCCTGACAGCTT
          TCCTGCCATGTGGCACACACTTTGGGAGCAGAGCCATAGCCCAAAGTGGACCGCCCTTGA
          GCTAGAAGTGTTGACTCAGGCGTGGGAAGGTGTAGAGCAGGCGGGTCACGGTGAGGAAGG
          AGTGGGGGGCTCAGTTGTCATGGGAGGTGCATGAATTCGTACTGCAGAGTGGCTGCTCAG
          GGGTCTCCTGTGTTGACATGTTATGTCAGGTTAAGCCATTTTAGCATTCTTAGTTTTCTG

8680      CTTGAGACCCACATGTGAAAACTGTCAGTCTGTTATCACTTAAGACAGAAGAAAATTGCC
          CTTGACTCTGGGCTGGCAGCAGGTGGAGACAAGGCCTGACAGCTTTCCTGCCATGTGGCA
          CACACTTTGGGAGCAGAGCCATAGCCCAAAGTGGACCGCCCTTGAGCTAGAAGTGTTGAC
          TCAGGCGTGGGAAGGTGTAGAGCAGGCGGGTCACGGTGAGGAAGGAGTGGGGGGCTCAGT
          TGTCATGGGAGGTGCATGAATTCGTACTGCAGAGTGGCTGCTCAGGGGTCTCCTGTGTTG
          [A,G]
          CATGTTATGTCAGGTTAAGCCATTTTAGCATTCTTAGTTTTCTGAGGAAACTCCACAGAA
          AGTTTTGCTTTATTTCTTAGAAGTAAGGACAGATACCGGTTTCTCACCTGTCCTCTGCTC
          CTGTAGGCACATTGAGGATAAGTCCACCGTGTTTGGGACTGCGCTCAACTATGTGTCTCT
          CAGAATTCTGGGTGTTGGGCCTGACGATCCTGACCTGGTACGAGCCCGGAACATTCTTCA
          CAAGAAAGGTACGGCATGTGCAGCATGTGCTGGGCCAGGGGTTCGTGTCAACTCGATAAT

8700      ACTGTCAGTCTGTTATCACTTAAGACAGAAGAAAATTGCCCTTGACTCTGGGCTGGCAGC
          AGGTGGAGACAAGGCCTGACAGCTTTCCTGCCATGTGGCACACACTTTGGGAGCAGAGCC
          ATAGCCCAAAGTGGACCGCCCTTGAGCTAGAAGTGTTGACTCAGGCGTGGGAAGGTGTAG
          AGCAGGCGGGTCACGGTGAGGAAGGAGTGGGGGGCTCAGTTGTCATGGGAGGTGCATGAA
          TTCGTACTGCAGAGTGGCTGCTCAGGGGTCTCCTGTGTTGACATGTTATGTCAGGTTAAG
          [C,G]
          CATTTTAGCATTCTTAGTTTTCTGAGGAAACTCCACAGAAAGTTTTGCTTTATTTCTTAG
          AAGTAAGGACAGATACCGGTTTCTCACCTGTCCTCTGCTCCTGTAGGCACATTGAGGATA
          AGTCCACCGTGTTTGGGACTGCGCTCAACTATGTGTCTCTCAGAATTCTGGGTGTTGGGC
          CTGACGATCCTGACCTGGTACGAGCCCGGAACATTCTTCACAAGAAAGGTACGGCATGTG
          CAGCATGTGCTGGGCCAGGGGTTCGTGTCAACTCGATAATGAGCTCTCACAAACGAGATA

8996      TAAGCCATTTTAGCATTCTTAGTTTTCTGAGGAAACTCCACAGAAAGTTTTGCTTTATTT
          CTTAGAAGTAAGGACAGATACCGGTTTCTCACCTGTCCTCTGCTCCTGTAGGCACATTGA
          GGATAAGTCCACCGTGTTTGGGACTGCGCTCAACTATGTGTCTCTCAGAATTCTGGGTGT
          TGGGCCTGACGATCCTGACCTGGTACGAGCCCGGAACATTCTTCACAAGAAAGGTACGGC
          ATGTGCAGCATGTGCTGGGCCAGGGGTTCGTGTCAACTCGATAATGAGCTCTCACAAACG
          [A,C]
          GATACAGAAAGATGCACTTGCAGCTGAAACAGTGGGCAAAAGCACATGAGCAGGGAATTT
          GTCAAAGCAGAAGTAGGCAGACACTGTTTAACCTAGGCATCATTTTTTAAAAAAGCAAAT
          TAAGAGCCAGGCACAGTGAGTGGCTCACGCCTGCAATTCCAGCACTTTGGGAGACTGAGG
          TAGAAGGACCACTTCAACCTAAGAGTTCGAGGCCAGCCTGGGCAACATAGTGAGACCTGG
          TCTCTACAAAAACAATAAAATATTAGCCAGGTGTGATGATATGCACCTGTAGTCTCAGCT

10590     CATGAGATCCTGCCTTCTTTCTTGGTGAGCTTGTCACTATTGTCCTCAGTTCACTGTCAG
          CCTTTGGTGTCGTTGATGCTGCGTCCCCAAGGCTGCTGTCCGGTTCCCACCACACTCCTG
          GCGCCTGCCTGGTGAAGGAACGTGTTTAGGCTGCACTTTGCCTAGTAGCTTTGTGGGTCT
          TTATTGACTTTTGCATACCTTTTGGGGTTTGGAGCAGGGACTCCTCAGAAGCATGTTTAG
          ATGGTGTGGCTGTGCCAGGACTGCTGCTGCTGAAGTGGCTCTGGCATGGGGCCAGCGTGC
          [T,C]
          GGAGCTACTCTGGAGTCTAGGGTCGTCTTTGTTCCCATACAGGACCAGTCTGCCAAGTGG
          AGATGACACAGACTGGGGCAGCTCAGGCTTGGCTCAGAGGGCGAGGCTGAGTGTGCGCTG
          TCACTTCCCCACCTTGCCTTCTCCAGGCGCATGTGCACCTGGGCCCCTCGCTCACCTGAG
          CACTGAGGTGTCCCTGGACCTTCCCAGGTAGCTGTCTTCATGTGCTCCTTCCTGGGGCCA
          GGGGTTGCAAACACCTCTCCTGGGGCTGGACACACACACTCCCAGGAAAGCCACTGGTTC
```

FIGURE 3R

11090   CTTCCCAGGTAGCTGTCTTCATGTGCTCCTTCCTGGGGCCAGGGGTTGCAAACACCTCTC
CTGGGGCTGGACACACACACTCCCAGGAAAGCCACTGGTTCCACCTAGGGGGCCGTGTAT
CCAGGCAAGTTCTCAGCACTCTGGAACCTGCTTCGCACATGGGGGTCGCAAGATCCACAT
GAGGCTGCCCTTGCCTCATGGAGAGGGGCACACGTGACTCCCAGAGGGTGAAGCTTCCCA
GCTAGAGGCAGTGCAGACTTTGCTGACAGGAAGCAGATGACGTGGGCCTATTCTCTCCCC
[G,C]
CTCAGGTGGTGCTGTGGCCATCCCCTCCTGGGGGAAGTTCTGGCTGGCTGTCCTGAATGT
TTACAGCTGGGAAGGCCTCAATACCCTGTTCCCAGAGATGTGGTATGTCTGCTGTTGATT
GGGTTGTTGGGTCGCTGCTGCTGTCCCGGGGAGTAGAGTGACAGGGACCGTGGGTCAGGT
GCAGGCTGTGACAGCAGAGAGGGGTGGGCATTCTGTGGGTGGGTGGAGTTAGGCTCCTGG
CAGAGGCCCTGATCAAGCTTGAGTCCTGTAGGGGTACAGAAAGGGGGAGGTTCCCAATTG

11710   CATGGATGGAGGTACCCCGAGTCAGGCTGCAGGCAGGGCTGGGTGGCTTCCCTCTTGCTG
TGGAAGACTCAGCATCTGTAGAAGTGGGGGGGTGCCCCTCCCCCAGCCTGCACAGGGGCG
TCCTGTGTTGCTGCTGCTGCGTTTGTCTCCTTTGCTGGTGAATGTGAAGTGTGTCCCGAC
GTGACACCTCACCTGTGGACTCAGCGTGTGTGCCTTTAAAAGATCAGTGTCTGTGGCCAG
GTGGGGTGGCTCATGCCTGTAATCCCAGCACTTCGGGAGGCCGAGGCGGGCAGATCACGA
[G,A]
GTCAAGGGATCGAGACCATCCTGGCCAACATAGTGAAATCCCGTCTCTACTAAAAATACA
AAAATTAGCTGGGCGTGGCGGCGCGTGCCTCTAGTTTCCAGCTACTCGGGAGGCTGAGGC
AGGAGAATCACTTGACCCTGGGAGGCAGAGGTTACCGTGAGCCGAGATCGTGCCACCATA
TTCCAGCCTGGCGACGGAGTGAGACTCTGTCTCAAAAAAAAAAAAAAAAAAGATCAGTGTT
TGTTTTTTTAAACAGAACCACATACTGTTTAAATACCCAGCAAAATCAACATTAATTTCT

12591   GTGCTGGGAGCCATGAGCCACTGCTCCCGGCCTTATGTGGTGTCTTTAACCAGTGTCTTG
TAACATTTTATGGCTATCTATTGAAAGCAGTGGACATCTCCCCAGAAAACACTCGTGCAT
ATGAGTTTACCCCGTTATGCATTTTGGGAAGTGAGACCCTGGAACCACACAGAGCCCCTG
CTGGCTTCCTTGAGTGTTGTGGGAACCCTGGTGGGGGTGTCCCCTACAGAGCTATCATCA
GGGCTGGGGGGGTCCCTTGTGTTAGATGACTTTGGTGCGGGGGTGGGGGGTGGGGGGTCA
[G,A]
GTTAGGGGAGGCAGGAAGTGAAGGGGCCGCTCAAGAAAGGACAGCAGCAGTGTCCTGATG
CAAAGGCCGGGGGCTTAACCCCGGAAGCCAGTTTGGGTGGTGACGGGGAGGCACAGGGAT
GGTGAGATCACCCCGGGAGGGTAGACAGAGATACCAGAGTAGGGGGCAGGGTTAGGGTGC
CGCTACCTGAGGCGGGCCGTAGAGCACATAGGTTGGGAGGTGTCCTGGGGCCATTCAAAT
GCCCGCTGGACTCTGCGCCTCGCCCGTGTGTAATGAGCGGCAGAGGAAGGACTGAGACGG

13431   AAAAACACCCTTTTAAGTAAGTGGGTGTGAAAGTGGGCCAAGGCCTGATGCCACAGTCAG
GGAGCAGGGAAGGCTCAGCATTGCTCACCCTCACTTAAGGATGGGGCTAGCATCACATAA
GGCATCACATAAGGATGGGGGCTAGCAGGGAAAGGGAGAGAAAACACATGAGGCACACAC
AGACCCTGGGAAGCTGGTGGAGCTGTGCTAACGTCAGCAGACCAGTGATCAAAGACCCAG
GCCTTGGGGAGATTCCACAGACCTACAGACCTACAGTTTCTTTTTTCTTTCTTTTTCTTT
[-,T]
TTTTTTTTTTTTTTGAGACAGAGTCTCTCGCTCTGTCACCAGGCTGTGTGCAGTGGCACAA
TCTCGGCTCACTGCAACCTCCTCCCAGGTTCAAGCGATTCTCCTGCCTCAGCCTCCCGAG
TAGCTGGGACTAGAGGCACACACCACCATGCCTGGCTTATTTTTGTATTTTTAGTAGAGA
TGGGGTTTCGCCATGTTGGTCAGGCTGGTCTCAAACTCCTGACCTCAAGTGATCCACCAG
CCTCGGCCTCCCAAAGTGCTAGGGTTACAGGCGTGAGCCACCGTGCCCCTCCTAAAGTTT

14746   CATCCTATAATAAACAGTGAGCAAGCTCTGCCCAGAGGGGACTTGTGCTATGGGACAGTC
AGTAGCTGTAGCCCAGGGGTTCCTGGGGGGGACTTCCAGGACTCAAGGGATGCAGGAGGCA
GATGTGCACTGTGTCCTCTGGAAGCAGGCCTGAGGCGAGGTTTGAGGTGCAGGATGTTTA
TCAGGCCTGCCATGGGGAAGAAGGAGGGGCAGAGGGAGGAAATGAGCTTCTGGGCAGACC
TGGGACTCATGGAGCTGGGGAGCTCCTCAGAGCGGTCCTCCCATAGGGGGCCTTCATGTG
[C,G]
CCTCGGGGTCAGTTGCTGGAGGGACCCCCACCCAGGAAGGGACTGGCCCAGGGCCCTGAG
GGCGGATGGTGGGAGGCCACCCCTCCTGGTTTGAGCCAGGCCTACCAGGTGCTCCCAGGC
CCCAAGGCTCAGACACTGCCCCTACCAGGAGCTCTATGTGGAGGACTTCGCCAGCATTGA
CTGGCTGGCGCAGAGGAACAACGTGGCCCCCGACGAGCTGTACACGCCGCACAGCTGGCT
GCTCCGCGTGGTATATGGTGAGCGCCTCCTGAGGGGCCGGCAGGGCAGCCCAGGGTCAGG

14975   CTGGGCAGACCTGGGACTCATGGAGCTGGGGAGCTCCTCAGAGCGGTCCTCCCATAGGGG
GCCTTCATGTGCCCTCGGGGTCAGTTGCTGGAGGGACCCCCACCCAGGAAGGGACTGGCC
CAGGGCCCTGAGGGCGGATGGTGGGAGGCCACCCCTCCTGGTTTGAGCCAGGCCTACCAG

FIGURE 3S

```
GTGCTCCCAGGCCCCAAGGCTCAGACACTGCCCCTACCAGGAGCTCTATGTGGAGGACTT
CGCCAGCATTGACTGGCTGGCGCAGAGGAACAACGTGGCCCCCGACGAGCTGTACACGCC
[G,C]
CACAGCTGGCTGCTCCGCGTGGTATATGGTGAGCGCCTCCTGAGGGGCCGGCAGGGCAGC
CCAGGGTCAGGGTCAGGGTGTCGCCCACTCATTCACGCACTCATCCCCTGCCAGCGGCAC
TGGGCCACCTCCTCTGTGCCAGGCCCCAGGGGGCGGGATCTCATCGCCCTGCCCCTCCAC
CCTGAGAACCAGCTGGTCTTCTACTCTCAGGAGTCCACCCTGTGCAAGGGTGTGTGGTAG
GAGGTGTGGGGCAGCCCCTCCTGGGCAGGGAAGGAGGAGCTCAGAGACCAGGCCTGGGGG

16031    TTCCGTCCCCCACCGGCTCTTGTCCTCAGTGTGCCTGGACACTCTCCTAGAGGCCCCTCC
         CTGAGATCTTGCTGGCTAGCTGGCTAGCTGGGAGGGGTGCTTTTTCCTCACTTGGTTCCC
         TCTCCCCAAACAGTTCATCATTCGCCATTCTCCCGTGGGGTTTAGACATGCCCAGGGTGG
         GTGGGAGTAGCAGGTGCCACTCCTGATTCCTCCTGCCTAGCTAGGGACTTGGAGCTCTCA
         CCTCTGTGGGGCCTGCAGGGGTCCAGGTGTGGCCAGTTCAGTGACCTTAGAGGGTGCAAT
         [C,T]
         CCCGGGCTGTGCTGGTGCGTGGCCGCCTCCTGACAGAGTCAGCAGGCCCTGGGCTGTGCT
         GCAGCTGCTGCCGTAGCTGTGCGCGTAGCTGCTGCGGTGTAGTGGGTTGGCTTAGGCATT
         CTCTGGACATACCCAGGTGGCACTGGGCCACTGAGTCCCACCCTGACACTGCATCTCGGA
         TTTTCTGGGCCTCATGCCACCTCAGTGGATCACAAATCCTGACTGACCCTGCAGCGGGTC
         CCTTGTTTTTTGCTCAGCAGTGATGTGGTTCTTTGTGGGTTTTGGTTTAATCCCATATAG

16891    GTGGAGTGCTGCTCCTTCTCACAGCCTAAGGCAGGCTGTGGCCTTGGCCGACACTGCCTC
         TGTCTGAGTTGGGTCCTGGGGACACAGTTGTTGCCCATCCTCGCTCAGGAAATGCCTGTT
         AGAGCAGAAGGCCCCTGTCCTGGCCCTGAGTGATCTGCACGGCACTTTATGCCTGGGGGC
         TGCTGTGGATCTGGACGAGACCTTGTCCCTGGAGGCTGCTGTGGGTCTGGAGCGGAGCCT
         TGACAGGGCTGTCTCTCCTGCAGATCTCGAAAACCATCAACATGCTTGTGCGCTGGTATG
         [-,T]
         GGACGGGCCCGCCTCCACTGCCTTCCAGGAGCATGTCTCCAGAATCCCGGACTATCTCTG
         GTGAGTGTGGCTGGGATATGCTGGCGGGGCCTCTCACGAAGACTGGATCTGAGCCCCAGC
         TGCATCCCAGTGAGGGGGCCCCCACGGTGCCATCTGGGAATACTGCCAGGGAATACCTCC
         AGGAACCAGCAGTGTCAGGGCTTGTGGAAGCCACTGAGGGTTGTCTTTGAATTGGAAGAT
         TTGCCACCCAGTGGAAGTGTGGGGGTGTTCCCAGAAGGTAGAGTGAGGAAGGGGGTGGTAG

19359    CCACACACCACCCCCTGCCCAGTCCCCATGTCTGTCTGGTCAGTGCCCAGCTCTGTCTCA
         CTAGGGTTTGGTCACCGGCCCTTTGAACTGAGACCAGGCTGTGTACCTGTGAGCCCAGCT
         CGGGGTGAGATTTGAGGTGGAGCCTTCCCAGCCCTGTGCAGAATTCCCATCACCTCCAGG
         TGTACTCAGAAATGGGGATCATTGGCCAGGTGCGGTGGCTCACGCCTGTAATCCCTACAC
         TTTGGGAGGCCAAGGTGGGCGGATCACAAGGTCAGGAGATAGAGACCATCCTGGCTAACA
         [C,T]
         GGTGAAACCCCGATGCTACTAAAAAATACAAAAAAAATTAGCTGGATGTGCTGGCAGGAG
         CCTGTAATCCCAGCTACTCCGGAGGCTGAGGCAGGAGAATGGCGTGAACCCAGGAGGCGG
         AGCTTGCAGCGAGCTGAGATCACGCCACTGCACTCCAGCCTGGGCAACAGAGCGAGACTT
         CATCTCAAAAAAAAAAGAAATGGGGTCATTTCCAGGCATCACCATGACTGAGGTGCGCCA
         CTGTCATTGGGTGAGAGCAGCTGGATGCTCTATGTGTAGGTGCTGGAGCCTCTGAGGGAT

19405    CCAGCTCTGTCTCACTAGGGTTTGGTCACCGGCCCTTTGAACTGAGACCAGGCTGTGTAC
         CTGTGAGCCCAGCTCGGGGTGAGATTTGAGGTGGAGCCTTCCCAGCCCTGTGCAGAATTC
         CCATCACCTCCAGGTGTACTCAGAAATGGGGATCATTGGCCAGGTGCGGTGGCTCACGCC
         TGTAATCCCTACACTTTGGGAGGCCAAGGTGGGCGGATCACAAGGTCAGGAGATAGAGAC
         CATCCTGGCTAACACGGTGAAACCCCGATGCTACTAAAAAATACAAAAAAAATTAGCTGG
         [A,G]
         TGTGCTGGCAGGAGCCTGTAATCCCAGCTACTCCGGAGGCTGAGGCAGGAGAATGGCGTG
         AACCCAGGAGGCGGAGCTTGCAGCGAGCTGAGATCACGCCACTGCACTCCAGCCTGGGCA
         ACAGAGCGAGACTTCATCTCAAAAAAAAAAGAAATGGGGTCATTTCCAGGCATCACCATG
         ACTGAGGTGCGCCACTGTCATTGGGTGAGAGCAGCTGGATGCTCTATGTGTAGGTGCTGG
         AGCCTCTGAGGGATCGTCCAGTCCTAGAAGTGTCCTCAGAGGGACACTGTCCTGCCTGGT

19653    CTAACACGGTGAAACCCCGATGCTACTAAAAAATACAAAAAAAATTAGCTGGATGTGCTG
         GCAGGAGCCTGTAATCCCAGCTACTCCGGAGGCTGAGGCAGGAGAATGGCGTGAACCCAG
         GAGGCGGAGCTTGCAGCGAGCTGAGATCACGCCACTGCACTCCAGCCTGGGCAACAGAGC
         GAGACTTCATCTCAAAAAAAAAAGAAATGGGGTCATTTCCAGGCATCACCATGACTGAGG
         TGCGCCACTGTCATTGGGTGAGAGCAGCTGGATGCTCTATGTGTAGGTGCTGGAGCCTCT
         [G,A]
```

FIGURE 3T

AGGGATCGTCCAGTCCTAGAAGTGTCCTCAGAGGGACACTGTCCTGCCTGGTGGCCCATG
AAGAAAGGGAGGGCTCCCTGAGTCTCCCTGACGTGTGTCTGCCTGCAGGGCTCAGCCTTC
TCTGAGGCCCTTGTCAGCCATGAGGGGTGCCCAGGGCTCAGAGCCTGAGGCTGAGCGTTG
GCTGGGTGGGAGCCCCCACACCTGGCCCTCAGGCGCCCATTGGATCCTGGAGGCAGTGGC
TGGGAGTGGGAGGGGCTGCATCTGCTGCTGTAACACCATCCTTTGTGTGTAGGGCACCAA

19742   AGGCTGAGGCAGGAGAATGGCGTGAACCCAGGAGGCGGAGCTTGCAGCGAGCTGAGATCA
CGCCACTGCACTCCAGCCTGGGCAACAGAGCGAGACTTCATCTCAAAAAAAAAAAGAAATG
GGGTCATTTCCAGGCATCACCATGACTGAGGTGCGCCACTGTCATTGGGTGAGAGCAGCT
GGATGCTCTATGTGTAGGTGCTGGAGCCTCTGAGGGATCGTCCAGTCCTAGAAGTGTCCT
CAGAGGGACACTGTCCTGCCTGGTGGCCCATGAAGAAAGGGAGGGCTCCCTGAGTCTCCC
[T,C]
GACGTGTGTCTGCCTGCAGGGCTCAGCCTTCTCTGAGGCCCTTGTCAGCCATGAGGGGTG
CCCAGGGCTCAGAGCCTGAGGCTGAGCGTTGGCTGGGTGGGAGCCCCCACACCTGGCCCT
CAGGCGCCCATTGGATCCTGGAGGCAGTGGCTGGGAGTGGGAGGGGCTGCATCTGCTGCT
GTAACACCATCCTTTGTGTGTAGGGCACCAACGGCTCACAGATCTGGGACACCGCATTCG
CCATCCAGGCTCTGCTTGAGGTTCGTGGCTCCTTCTCTTTTCTCAGCCTCAGCTGACCTT

20054   GCCTGCAGGGCTCAGCCTTCTCTGAGGCCCTTGTCAGCCATGAGGGGTGCCCAGGGCTCA
GAGCCTGAGGCTGAGCGTTGGCTGGGTGGGAGCCCCCACACCTGGCCCTCAGGCGCCCAT
TGGATCCTGGAGGCAGTGGCTGGGAGTGGGAGGGGCTGCATCTGCTGCTGTAACACCATC
CTTTGTGTGTAGGGCACCAACGGCTCACAGATCTGGGACACCGCATTCGCCATCCAGGCT
CTGCTTGAGGTTCGTGGCTCCTTCTCTTTTCTCAGCCTCAGCTGACCTTCCTGTGCACGT
[A,G]
AGCCCACGCATCCACCTGAGGGCAGCACTGCTGGCCACACACTTGCCACTCCTCGATACT
TCCAGTGACCTGGGCTCTGGCCTCTGGCTTCAGAGGGTCGTGCTGTGGAGGGGGCGGCCT
TGGCCAGCAGCCTTGGGTGTTGGGCTGGGTCGGGGCCTTGGGAGGGCAGGGGCTGGAGG
CTGTGTGAGAAGGGGAGTCTGGTGAAGGCTGTTTCTGAGAGTGCAGGCAGGAGTGGGACT
CCAGGCTCTTCTTAGAACTGGAACTGCTTGGGCCAGGCACGGTGGCTCACACCTGTAATC

20627   CCAGGCACGGTGGCTCACACCTGTAATCCCAGCACTTTGGGAGGCCGAGGAGGGTGGATC
ACGAGGTCAGGAGTTCAAGACCAGCCTGGCCAAGATGGTGAAACCCCGTCTCTACTAAAA
GTACACAAAAATTAGCCAAGCGTGGTGGCGGGCACCTGTAATCCCAGCTACTTGGGAGGC
TGAGGCAGAGAATTGCTTGAACCCGGGAAGTGGAGGGTGCAGCGAGCCGAGATTGTGCCA
CTGCACTCCAGCCTGGGTGACAGAGAGAGGCTCCGTCTCAAAAAAAAAAAAAAAAAAAAAA
[-,A,G]
AACTGGAACTGTTTGTTATGGGCATTCTCGAGCCAGTACTGGAGAAAAACGAGAGTGGAT
TTTTATGCCGGTGGGAATGAGGTAGGTGGGATTCTGAAGGTGTTTCTGGAGAGCCCTGAG
GGCTGGGCCACGCAAAGGGCCTGCCTACACAGGGTGCTGGAGACCCTCTGGGCATGGATG
CTGGCCAGGCAGGGGGGTGCTGGCATCCATAAATGGTCTCCTGCGCCCTTCCATCTTCAG
TCATATCTCATGGACTTTTGCTGTTTTGTCTTTAAAGGTAAGTGCAGCAGGAGACCCTGG

21337   AGCCTCTCTGTCCTGCTGTCTCTTCCAGGCGGGCGGGCACCACAGGCCCGAGTTTTCGTC
CTGCCTGCAGAAGGCTCATGAGTTCCTGAGGCTCTCACAGGTGAGGCCGGTGCCTGGGGC
TCTGAGGGGGCTGAAGAGGGGGATCAGGGCTGGGAGCTCCTGCAGGCAGAAGTGCCCACC
TCACCTCCACCCTGCCCTATTTCCTGCACTGGTGTTTCAGGGTCACCCCCACCCTCCCAT
CCCCTCCCTAGCCCCTGCTCCATCCACCGGTCCTCCTCGGGCTGGCCTCACCTGGGGCAG
[T,C]
TCTCTGAGGCCTGCAGGGTGCTGGGGGTGCTGGCAGTTTCTGCGTCCTGCTCATGTTGGA
GCCACTGTGTGCAAGGGCCAGGCACGGGCAGGGCTGTGTACCCTGAGCTGCACAGCCTA
CACGGCACCTCCATGTCTCTGAAGCACCTTCTGCCCATGGAGGTGACGCCAGCCTGTGGA
CTTGCCCTCCTGAGACTGTTTGCAGCAAAAGCCCCGGTCCCTCCTGCCAGATCAGCTGCC
CACAGACCCTGCCCGAGCCCATAGTTTGACCTCAGTGTCTCTCACACGTGCCTGCACCCC

21894   GCCCATAGTTTGACCTCAGTGTCTCTCACACGTGCCTGCACCCCAGTCTGCAGCCACAGT
CATCCCATACATGCGCCCCAACCTCCCGTGTCTCCCACACCCTGTCCCGGCCACGGCCTC
AGCCAGTGTCCCTCTGCCTGGAACCGCTGCCCCCCAGCCCCGTCTCCCTCCCTTCAGCTC
TCACTAGGACATTGTTCTGCAGGGCTTCTGGGTCTTCCTGGCCTCTGTGTGGCCAAGGCT
GGCACCCATCTTGGGCTCAAGCAGAGGAGGGGCATTGTCCTGCTGTGCCTGGCCCAATGG
[C,T]
GGCCTGCTCCTGCTCCTGCCTCCTGCCCAGGACTTGCTCTGGGTGATGGGGACTTGGGGA
GGCTGACTGAACCCTACGGCACTCCAGGCCTCTTCCCTTCTCACTGAGGTGAGAGAGGCA
GCCAGAAGCTGAGGTTGTTCAGGAGGCATTGGGGGCGCCTGGCACAGAGCACACCCGCAG

FIGURE 3U

```
        AGACCTGGGCCCCCTCCCTGCCTTCTGGCCGGTGGGGAGATCACAGGGGAGTCAGGTGCT
        GACTCCCAGTCCCGTCTGGGCTGGTTTGAGCCCTCGCTGGCCAGTCACGTTTCCCAGCAG

23360   TGAGAACTGGGGTGTGGACACCCCCAGCCTGGAGTCATGGCTTGTGCTCTGCAGGGTGGC
        TTCTCCTTCAGTACGCTGGACTGCGGCTGGATCGTTTCTGACTGCACGGCTGAGGCCTTG
        AAGGCTGTGCTGCTCCTGCAGGAGAAGTGTCCCCATGTCACCGAGCACATCCCCAGAGAA
        CGGCTCTGCGATGCTGTGGCTGTGGTAAGGCTGTGGTCCCAGCAGCCCCGTCCATACCTC
        GTGTCCTGCAGATGAGCTGCGTGCTCACTTCCACTCCTGTGGGCTCCAGCCCAGCACACA
        [G,T]
        TCCGGCCAGGCCGTAGGAGCTTGTCCTTGGATGGTGTCTATATGTGGAGAACTGTGAGCT
        CTGGCTGGACCCCTAGGGGCCTTGCTGGGCTGTGTGCACAGGGCCCTGCACTGCGGAGCT
        GGTGTCCAGCCCAGCCACCGATACTTGGGGGAGCCGGCGTGGCCCCCAAGGTTTCTCTCT
        GGTGGTTTCCACTGGGTGTCTGAAGAGGGAATTTGTTGGTGTTGGTTTTGGTGCCACATC
        CTTTCAGCACATCTGGCTTTTGTGTGTGTGTTTCCCAGTGGAGACCCTGCCCTTTTCTGGCA

26758   CAACCTCCGTTTCCGGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCCGAGTAGCTGGGAC
        TACAGGTGCTGACCACCATGACTGGCCAATTTTTTGGTATTTTTATTAGAGACAGGGTTT
        TACCATGTTGTCCAAGCTGGTCTTGAACTCCTGACCTCAGGTGATCCACCTTCCTTGGCC
        TCCCAGAGTGCTGGGATTACAGGTGTGAGCCACCACACTGGCCTTTGCTATTTTCTTTCT
        CCTTTATTTTTCTAACTTGAATACTTAGATATTTGATTTTCAGGCTTTTATTGAAATATG
        [A,C]
        ATTTGAGGCTATAAATGAGTTTTGAGATATCATTCAGTTAAATGTGTGTTCTGGTGCTTG
        CTGTGGTAGCACAGATACTAAAAGTGTTTTCTGTTTCTACTGTTCTTCTCTGGCCCATGA
        GTTATGTGGGAGTATGCTGCTTCATTTACAATCTGAGAATGTTCTGGTGTGGTTTTTTTG
        GAAGCCGTGGATGGAGCAGGGGTTTTCTTGTGCTTCACAGGTGCAGCTAGGAGGGCACTG
        TGTCCAGGGTCTTCTGTCGGCCTGGCGTGGCCCTTGGCCATGTGCTGCTCTGCGGCATGA

27033   ATTTTCAGGCTTTTATTGAAATATGAATTTGAGGCTATAAATGAGTTTTGAGATATCATT
        CAGTTAAATGTGTGTTCTGGTGCTTGCTGTGGTAGCACAGATACTAAAAGTGTTTTCTGT
        TTCTACTGTTCTTCTCTGGCCCATGAGTTATGTGGGAGTATGCTGCTTCATTTACAATCT
        GAGAATGTTCTGGTGTGGTTTTTTTGGAAGCCGTGGATGGAGCAGGGGTTTTCTTGTGCT
        TCACAGGTGCAGCTAGGAGGGCACTGTGTCCAGGGTCTTCTGTCGGCCTGGCGTGGCCCT
        [T,C]
        GGCCATGTGCTGCTCTGCGGCATGAGGTGGGCGTGAGTTGTCCTCAGCCACATTTAGAGA
        ATTGGCCTTTTAAAAAATAGATCATCTTTTAAAAATCACTGTAATAAAAGTAAAGCAGGT
        TCTTTGCAAACAAGACTTGCAAAATACAGAGAAGCGCAAAGAAGAAGCTAAGTCGCCCCT
        CCTCGCCCCTGAAGGAGAATCTGCTGTTGCTGTTTGGTCTCCACATTTCCATGGCGGCTT
        GCTGCCCCTTTCACGCCTGGCCCACTTTGTGCCTGGTGAGGTTTCTAAAAGCCCCACCCT

27332   TTGGCCATGTGCTGCTCTGCGGCATGAGGTGGGCGTGAGTTGTCCTCAGCCACATTTAGA
        GAATTGGCCTTTTAAAAAATAGATCATCTTTTAAAAATCACTGTAATAAAAGTAAAGCAG
        GTTCTTTGCAAACAAGACTTGCAAAATACAGAGAAGCGCAAAGAAGAAGCTAAGTCGCCC
        CTCCTCGCCCCTGAAGGAGAATCTGCTGTTGCTGTTTGGTCTCCACATTTCCATGGCGGC
        TTGCTGCCCCTTTCACGCCTGGCCCACTTTGTGCCTGGTGAGGTTTCTAAAAGCCCCACC
        [C,A]
        TTGAGCGCGCTCCTCCAGCACGAGCAGTAATGGCACAGGTGTTGTGTCATTTTACTCAGT
        AGCCTCTGGGTTATTTTTCAGTTTTCCTTGTTGTTTTTTAGCTTTTTCCCCATTTTAACCT
        TAACTGGTATTTTCTTGTTAAATATTTATTCATGACCATTATTATTCCCTAGAGCCACAT
        GGCTTGGGGTCCACCTGCCTGGGTCCGCCCCCATCCCTGCCCCTTCTGGCTGTCTGACCT
        GGCCTGGTGACTTCTCTTCTCTGCTCATCTCTCTCCCTGCCTGAGTGGGCAAGAGTACAG

27538   TGTTGCTGTTTGGTCTCCACATTTCCATGGCGGCTTGCTGCCCCTTTCACGCCTGGCCCA
        CTTTGTGCCTGGTGAGGTTTCTAAAAGCCCCACCCTTGAGCGCGCTCCTCCAGCACGAGC
        AGTAATGGCACAGGTGTTGTGTCATTTTACTCAGTAGCCTCTGGGTTATTTTTCAGTTTT
        CCTTGTTGTTTTTTAGCTTTTTCCCCATTTTAACCTTAACTGGTATTTTCTTGTTAAATAT
        TTATTCATGACCATTATTATTCCCTAGAGCCACATGGCTTGGGGTCCACCTGCCTGGGTC
        [C,A]
        GCCCCCATCCCTGCCCCTTCTGGCTGTCTGACCTGGCCTGGTGACTTCTCTTCTCTGCTC
        ATCTCTCTCCCTGCCTGAGTGGGCAAGAGTACAGCCTCACAGAGTGGTGGGATTGTGTGA
        GATGCCACAGGGAAGCACATGTCAGTTGTTGTCACTGTGTAGAACAATGAGTCCCGGATG
        TGGCCCGCAGGGGAGCAATGGTGACTTAATCGCGGGCTTCCTCTGCATTTCTTTGGTGAC
        TTCCAAGCTAGAACATTCTTTTTTTGTTTATTTGTTTGAAGCAGGGTCTCACTCTGTTAC
```

FIGURE 3V

27625   CCCCACCCCTTGAGCGCGCTCCTCCAGCACGAGCAGTAATGGCACAGGTGTTGTGTCATTT
TACTCAGTAGCCTCTGGGTTATTTTTCAGTTTTCCTTGTTGTTTTTTAGCTTTTCCCCAT
TTTAACCTTAACTGGTATTTTCTTGTTAAATATTTATTCATGACCATTATTATTCCCTAG
AGCCACATGGCTTGGGGTCCACCTGCCTGGGTCCGCCCCCATCCCTGCCCCTTCTGGCTG
TCTGACCTGGCCTGGTGACTTCTCTTCTCTGCTCATCTCTCTCCCTGCCTGAGTGGGCAA
[G,C]
AGTACAGCCTCACAGAGTGGTGGGATTGTGTGAGATGCCACAGGGAAGCACATGTCAGTT
GTTGTCACTGTGTAGAACAATGAGTCCCGGATGTGGCCCGCAGGGGAGCAATGGTGACTT
AATCGCGGGCTTCCTCTGCATTTCTTTGGTGACTTCCAAGCTAGAACATTCTTTTTTTGT
TTATTTGTTTGAAGCAGGGTCTCACTCTGTTACCTAGGCTGGAGTGCAGTAGCAAAATCA
TGGCTCACCACAGTCTCAAACTTCCGGGCTCAAGCAATCCTCCCACCTCAGCCTCCTGAG

27736   TTTCCCCATTTTAACCTTAACTGGTATTTTCTTGTTAAATATTTATTCATGACCATTATT
ATTCCCTAGAGCCACATGGCTTGGGGTCCACCTGCCTGGGTCCGCCCCCATCCCTGCCCC
TTCTGGCTGTCTGACCTGGCCTGGTGACTTCTCTTCTCTGCTCATCTCTCTCCCTGCCTG
AGTGGGCAAGAGTACAGCCTCACAGAGTGGTGGGATTGTGTGAGATGCCACAGGGAAGCA
CATGTCAGTTGTTGTCACTGTGTAGAACAATGAGTCCCGGATGTGGCCCGCAGGGGAGCA
[A,G]
TGGTGACTTAATCGCGGGCTTCCTCTGCATTTCTTTGGTGACTTCCAAGCTAGAACATTC
TTTTTTTGTTTATTTGTTTGAAGCAGGGTCTCACTCTGTTACCTAGGCTGGAGTGCAGTA
GCAAAATCATGGCTCACCACAGTCTCAAACTTCCGGGCTCAAGCAATCCTCCCACCTCAG
CCTCCTGAGTAGCTGGGACTACAGGTGCATACCATCACCTGTGGCTAATTTTTTAAATGT
TTTGTATTTTTTAAATGTTGCTCAGGCTGGTCTTGAACTGCTGGGCTCAAGCAATCCTCC

30688   TACGCAATTGATTTTGATACTGATCTCATAGCTAGACAATTTTGCTAAACTTTTAAAAAA
ATTTATGTACTTTATCTTTTATAGCAGCTTTAAATTTACAGAAAATTTGAGTGGAAGATG
CAGTGTTCCCATAAAGCCGCTAACTCCTCGCACCTTCCCTCAAGTTTCCCCAGTACTAAC
ATCTTGCATTCAAGTGGTGCGTTTGCAACATTCATAAATTATTATCGTCCAGAGTCCATT
GTTTACATTCAGCTTCCTCTTCATGTTGTTCATTCTGTGGTTTCACAGATGTGTGATGCA
[T,C]
GTGCCCACCACTGCAGTGTCACACAGGATCTCACTGCCCCGGAGTCCTCTGCGCTGTCCC
CGCCTCCAGAACCCCTTAGTAGCAAACACTGATATTTTTACTGTCTCCATAGTTTTGCCT
TTTCAGACTGACCTATTTCACTTAGTAAGAAGCATTTAAGATTCCTGAGTCTCTTTCTAT
GGCTCAATAGCACATTTCTTTTTAGTGCTGAATAATATTCCATTGTCTGGATGTACCACA
GTTTATTCATTCACCTACTAAGGTGAATGTCTTGCTTGCTTCCAAGTTTTGGCAACTATG

31172   TCAATAGCACATTTCTTTTTAGTGCTGAATAATATTCCATTGTCTGGATGTACCACAGTT
TATTCATTCACCTACTAAGGTGAATGTCTTGCTTGCTTCCAAGTTTTGGCAACTATGAAT
AAAGTTGCTATCAATGTTAGCGTGCACATAAGTTTTCAGCTCATTTGGGTAAATGCCAAG
AAGCATGATTGCGGGATCCTATGGTAAGAGTGTGTTTAGTTCTGTAAGAAGCTGCCAAAC
TGTATCTTAAGTGGCTGCACCATTTGCGTTTCCACCAGCAATGATGAGCGTTTTGTTGCT
[C,T]
CACATCCTCACCAGCATTTGCTGTTGTGTTTTGGGTTTTAGCCTTTCTAAGAGGTGTGTA
GTGGTATCTCCTTGTTTCAATTTGCAATTCCCTAATGACATTATGTTAAAATCTTGTCAT
ATAGTTATTTGCCATCTGTGTATCTTTTTCAGTGATGTGTCCTTTAAAGTCTTTGGCTCA
TTTTTAAATTAAATTTTCTTATTGTTGAGTTTTAGTTCTTCATATATTTTGGCTGCCAGT
CCTTTATCAGATATGTCTTTCGCAAATATTTTCTGCCTGTGTCTTGTCTTTTCATTCTAT

31433   ATTTGCGTTTCCACCAGCAATGATGAGCGTTTTGTTGCTCCACATCCTCACCAGCATTTG
CTGTTGTGTTTTGGGTTTTAGCCTTTCTAAGAGGTGTGTAGTGGTATCTCCTTGTTTCAA
TTTGCAATTCCCTAATGACATTATGTTAAAATCTTGTCATATAGTTATTTGCCATCTGTG
TATCTTTTTCAGTGATGTGTCCTTTAAAGTCTTTGGCTCATTTTTAAATTAAATTTTCTT
ATTGTTGAGTTTTAGTTCTTCATATATTTTGGCTGCCAGTCCTTTATCAGATATGTCTTT
[C,T]
GCAAATATTTTCTGCCTGTGTCTTGTCTTTTCATTCTATTAACAGTATCTTTTGCAGAGC
CAGTTTTCATTTCAAGGAAGTCCAGCTTATCAATGTTCTCTTTCATGTATCATGTTTTTG
GTGTTGTATCTAAAAAGTTACTGCCAAGCCCAAGGGTACCTAGATTTTTTCCTGTGTTAT
ATTCTAGGATTTTTAAAGTTTTGCATTTTACATCTAGGTCCATGATTCATTTGAGTTAA
CTTTTGTGAAGGGTTTATGGTTTGTGTCTAGATTTTTTTTTTTTTTTTTTTTGCATGT

32660   CTCCTGGGCTTAAGGAATCCTCCTGTCTCAGCCTCCTGAGCAGCTAGGACCACAGGCATG
TGCCACTACGTTCAGCTAATTTTTCAATTTTTTTGTAGAGATGGGATCTTGCTCTGTTGC
CCAGGCTGGTCTCAAACTCCCGTCTGCTTTGAGATGATTATATATTTGTGTCCTTTGTTA

```
          ATTTAGAGGATTATTATGGATTTTTCTAATGTTAAGACACCTTTGTATTTCTGAGATCGA
          CCTTAGTATTGGTCTATATTTAAGACAGTATTCAGTTTCTCAGTTGTTTTTTGTTTTTTG
          [G,T]
          TTTTTTTTTTTGAGACAGAGTCTCTGTCTCCCAGGCTGGAGTCCAGTGGCACAATCTCAG
          CTCACCGCAAGCTCTGCCTCCCGGATTCACGCCATTCTCCTGCCTCAGCCTCCCGAGTAG
          CTGGGACTACAGGCGCCTGTCATCATGCCCAGCTAATTTTTTGTATTTTTAGTAGAGACG
          GGGTTTCACCATGTTAGCCAGGGTGGTCTCAATCTCCTGACCTCGTGATCTGCCCACCTC
          GATCTCCCAAAGTGCTGGGATTACAAGGCGTGAGCCACTGCGCCCGGCAGCAGTTTCTCA

  32981   TCTCTGTCTCCCAGGCTGGAGTCCAGTGGCACAATCTCAGCTCACCGCAAGCTCTGCCTC
          CCGGATTCACGCCATTCTCCTGCCTCAGCCTCCCGAGTAGCTGGGACTACAGGCGCCTGT
          CATCATGCCCAGCTAATTTTTTGTATTTTTAGTAGAGACGGGGTTTCACCATGTTAGCCA
          GGGTGGTCTCAATCTCCTGACCTCGTGATCTGCCCACCTCGATCTCCCAAAGTGCTGGGA
          TTACAAGGCGTGAGCCACTGCGCCCGGCAGCAGTTTCTCAGTTTTAATTTGGAGTTTTGC
          [A,C]
          TCTGTGTTCATGAGTGAGCCTGAAATTTTCACTTTTCCATATCTTATTTCTCTGGGTTCC
          TAGAATGAGCTAGAGAGTGTTCCTCCTTTCTGTTCTCTGGAAGAGTTTGTGTGAGATTAG
          AATGAGTGTGTCTGATAATTTAGTTGCATTCATTTATAAAATTCCTAGGCCTAGAGTTTT
          TTTTCTGGGAAAAGTTTACATTTTGACTCATTTTTTTAGTAGTTTTAGGACTGTTTAGGT
          TCTCTATTTCTTGATTGAGCCAGTTTTGATAAGTTAATCTTTCTAATTTGTAGATATTTT

  33557   AATCTTTCTAATTTGTAGATATTTTTCTCTAAGTTTGCAAATGTAATACATAAAACTTTCT
          TGTCATTTCTCACCATATCTGTAGTTCTATCTTTTTATTGCTAATATTACTAATTTGTAC
          TTTGACTATTTGTATTTGTTACCTGTTGCCGAGTAACAATATTAGTACAAACCTAGTGGC
          TTAGAACAACACACATTGATTACTTCACCGTTTCTGTGTGTCAGAAGTCCAGGCGCGGCC
          TCGCAGGTCGTCCTCTGCCTCAGGGTCTCTCCGGGCTTCAGTCAGGGTGTTAGCCAGGAC
          [T,C]
          GGGGTCTCGCCTGAGCTTCCAGTGAGGAAGGATCTGCCTCTGAGCACACAGGGTCCTCGG
          CACGATCCCATTCCTCAGCTGGAAGCTGCCGACTGCCGTCTGCTGCGGGGCCTCTCTAGA
          TGGCATCTTCACAAAAGCGAGAAGGGAGAGTTGGTAGAGGGAGTCTGCTAGCACCATGGG
          AGTCGCGGTCACACAGACCTCGGTCCCAGGACCCGCACCCATCAACCCTGCCGTGATCTG
          CTGGTTAAAGACAAGTCCCACGTCCCACAGGGTGACACTGGAGTAGACACTTCGCTCTGG

  33652   TATTGCTAATATTACTAATTTGTACTTTGACTATTTGTATTTGTTACCTGTTGCCGAGTA
          ACAATATTAGTACAAACCTAGTGGCTTAGAACAACACACATTGATTACTTCACCGTTTCT
          GTGTGTCAGAAGTCCAGGCGCGGCCTCGCAGGTCGTCCTCTGCCTCAGGGTCTCTCCGGG
          CTTCAGTCAGGGTGTTAGCCAGGACCGGGGTCTCGCCTGAGCTTCCAGTGAGGAAGGATC
          TGCCTCTGAGCACACAGGGTCCTCGGCACGATCCCATTCCTCAGCTGGAAGCTGCCGACT
          [G,A]
          CCGTCTGCTGCGGGGCCTCTCTAGATGGCATCTTCACAAAAGCGAGAAGGGAGAGTTGGT
          AGAGGGAGTCTGCTAGCACCATGGGAGTCGCGGTCACACAGACCTCGGTCCCAGGACCCG
          CACCCATCAACCCTGCCGTGATCTGCTGGTTAAAGACAAGTCCCACGTCCCACAGGGTGA
          CACTGGAGTAGACACTTCGCTCTGGCCTTTTCAGAGAACTGGTTATTTTTTGGAAATATC
          AGTTAGATGTAGGATGGGTCTTGTCTTCTAAATCTATTGTTTTCTCTCTAATTGATTTTT

  34390   CTCTGCCAATCCGCTTCCCGCTCTGGTGTCCTGTGGTTGCTTCTTTTTAAAACCCTCATC
          GGTCTGTGTAAACTGTTTATTTTTATGTGGTTTTTAAGGGAGACCATTCTCATTCTTTTG
          AGACCCTGGAAAGGATGGAATTGGGATAGGTAAACTGCTGTTTTACCAGAATGTTCACTG
          GACCAATCTCGTGTTCCAGGGAGACCCTCACGCAGGGCTTAGAGTTCTGTCGGCGGCAGC
          AGAGGGCCGATGGCTCCTGGGAAGGGTGAGTGAGCCTCCACTCGTGAGTGCAGAGATGCA
          [T,C]
          GGGATCCAGAGGTTTCTGCTCTCACACACTGCGTTCATAAATGTTGGCTTGTATGTTGTT
          GCTACACCAGAAGTTTCTGGAAGTGAGCTGCCAGCCCGTGACTTCTGGGGGACCTCGTTC
          CTTTGTGGCATGCGTGGCCTTTGCCCCGGTGGAAATTGCTCAGTACGTTGCTGGGCGCAG
          CCGGGCTGCTGGGAGCGCGCTGTAGCCTGAGCGTGGCTATTCCCTCCACCCTTTCTGCTT
          GCTCTTAGGGTCCAGCAGACAGAGCTGCTGTCTTCCACGGCCTTAATGCCTGAGGCACTG

  34399   TCCGCTTCCCGCTCTGGTGTCCTGTGGTTGCTTCTTTTTAAAACCCTCATCGGTCTGTGT
          AAACTGTTTATTTTTATGTGGTTTTTAAGGGAGACCATTCTCATTCTTTTGAGACCCTGG
          AAAGGATGGAATTGGGATAGGTAAACTGCTGTTTTACCAGAATGTTCACTGGACCAATCT
          CGTGTTCCAGGGAGACCCTCACGCAGGGCTTAGAGTTCTGTCGGCGGCAGCAGAGGGCCG
          ATGGCTCCTGGGAAGGGTGAGTGAGCCTCCACTCGTGAGTGCAGAGATGCATGGGATCCA
          [G,C]
```

FIGURE 3X

AGGTTTCTGCTCTCACACACTGCGTTCATAAATGTTGGCTTGTATGTTGTTGCTACACCA
GAAGTTTCTGGAAGTGAGCTGCCAGCCCGTGACTTCTGGGGGACCTCGTTCCTTTGTGGC
ATGCGTGGCCTTTGCCCCGGTGGAAATTGCTCAGTACGTTGCTGGGCGCAGCCGGGCTGC
TGGGAGCGCGCTGTAGCCTGAGCGTGGCTATTCCCTCCACCCTTTCTGCTTGCTCTTAGG
GTCCAGCAGACAGAGCTGCTGTCTTCCACGGCCTTAATGCCTGAGGCACTGGAGTTGGTG

34989   TGGAGTTGGTGGGCTGGCTGGGGCACGTGTGATTGTTGCAGAATGCGTGTTGTTTCACAC
         ACCGGCTGTGAACAGGGTGGAAGGGCTGAGGCTCTCCCTGTTTCCCTCCAGCTCCTGGGG
         AGTTTGCTTCACCTACGGCACCTGGTTTGGCCTGGAGGCCTTCGCCTGTATGGGGCAGAC
         CTACCGAGATGGGTGAGTGAGTGCCTGTCCTCTGGTGGGTGGGGGTTCTCAACCCAATGC
         TCTGTCATGAGTGTTTTTTGCTTTGACATTTGGTTTTAGGGTTTGTTTGTTTGTTTGTTT
         [G,-]
         TTTTTGAGACGGAGTCTCGCTCTGTCAACCGGGCTGACATGCAGTGGCATGATCCTAGCT
         CACTGCAGTCTCAAACTCGTGGGCTCAAGCGATCCTCCCGAGTAGCTGGGATCACAGGTG
         CACGCCACCACCCCGGGCTAATCTTTTAAAACTTTTATGTAGAGATGGAGTCTTGCTGTG
         TTGCTCACACTGGTTTGGGCTCAAGCAGTCTTCCTACCTCGGCCTTCCAAAGTGCTGGGG
         TTACAGGCATGAGCCAATGTGCCTGGCCTGTTTTTAATATTTTTAAACAGTGAGATAAGA

35067   GGAAGGGCTGAGGCTCTCCCTGTTTCCCTCCAGCTCCTGGGGAGTTTGCTTCACCTACGG
         CACCTGGTTTGGCCTGGAGGCCTTCGCCTGTATGGGGCAGACCTACCGAGATGGGTGAGT
         GAGTGCCTGTCCTCTGGTGGGTGGGGGTTCTCAACCCAATGCTCTGTCATGAGTGTTTTT
         TGCTTTGACATTTGGTTTTAGGGTTTGTTTGTTTGTTTGTTTGTTTTTGAGACGGAGTCT
         CGCTCTGTCAACCGGGCTGACATGCAGTGGCATGATCCTAGCTCACTGCAGTCTCAAACT
         [C,G]
         GTGGGCTCAAGCGATCCTCCCGAGTAGCTGGGATCACAGGTGCACGCCACCACCCCGGGC
         TAATCTTTTAAAACTTTTATGTAGAGATGGAGTCTTGCTGTGTTGCTCACACTGGTTTGG
         GCTCAAGCAGTCTTCCTACCTCGGCCTTCCAAAGTGCTGGGGTTACAGGCATGAGCCAAT
         GTGCCTGGCCTGTTTTTAATATTTTTAAACAGTGAGATAAGATCCCCGGTTGAAATGAAG
         ATGTTTCCCTGGTCCCACAGCTCTCTGGAGCTTCCTGACATGTATGCTGGAGGGACGCTT

35495   CAGTCTTCCTACCTCGGCCTTCCAAAGTGCTGGGGTTACAGGCATGAGCCAATGTGCCTG
         GCCTGTTTTTAATATTTTTAAACAGTGAGATAAGATCCCCGGTTGAAATGAAGATGTTTC
         CCTGGTCCCACAGCTCTCTGGAGCTTCCTGACATGTATGCTGGAGGGACGCTTCTGGTCT
         CCGGCCCCTCCAGGCATACAGATGCCTCCCAACCCTGAGTAGGAAGATTAGGGTCCACGG
         CCTCGCTGGAGCGGGTTAGAAGGCAGGAGATCTCCGGTCCCAGCCGTGTCTCCAGCCGCC
         [G,A]
         GACTCTCTCCCAGCCCTGTCTCCAGCTGCCCCACTGTCTCCCAGAGTCTGCCGTGTGGAT
         GTTTAGAGGTGGGGAGCACCGTGCTTGGCTGAGTGCAGCTTGTGAGACGCTGCTCCCAAG
         CACTGCAGACCTCACTCAGCCTGACGCGTCCGTGAGGCCATCCTCGGTACTCGCATGTCC
         CTTTGTCTTCCCAGCGACTCTGGGAGGCAGGAGTATCTGTTCCCAGTTCACATCTGCAAA
         AGTCAAGCTCGGGTTTCAGTAGTGGCCCATGGCCCTTAGGTAGGGTGGCCCCATCGTGCA

36001   GGCAGGAGTATCTGTTCCCAGTTCACATCTGCAAAAGTCAAGCTCGGGTTTCAGTAGTGG
         CCCATGGCCCTTAGGTAGGGTGGCCCCATCGTGCAGGCTCCTCCCCGTACCCCAAGGCAG
         CCTGCTGGGGTGAGAAGCCAGGGGTCTGGGACCTTCCTTGGTGTGATGGTGTCTCCTGTC
         TCTGGTCTTTGCAGGACTGCCTGTGCAGAGGTCTCCCGGGCCTGTGACTTCCTGCTGTCC
         CGGCAGATGGCAGACGGAGGCTGGGGGGAGGACTTTGAGTCCTGCGAGGAGCGGCGTTAT
         [T,G]
         TGCAGAGTGCCCAGTCCCAGATCCATAACACATGCTGGGCCATGATGGGGCTGATGGCCG
         TTCGGTGGGGACGACGGGACCGTCCCTGAGCCTTGGGTTTGGGTAGAGGAGGGACACTCA
         GCTGTGAGCCGGTGGCCTGGGCTGAGTGAATGTAGAGAGGAGGGGAGGCCTGTGGGCCAG
         GTCAGCTGCCACTCTGGGAACAGACACCTACAAGAGCCACATGCCTGGTTCCTGGGGCAA
         GAACGTGGGCTGCTCTGACCAAGTGGGGCCCTGCAGAGAGGCTCGCCTCTTAGAAGTGAA

38948   ACAGCTCCCCAATGGCGACTGGCCGCAGGTATGCCGCCAGGGACCTGAGCGCACAAGGCC
         CAGCACTGACCTCCAGCGTGCATGGCTGTTTCCACGTCCCCCTGCTCTGTGTCCTTTTTG
         GGGTACTTTGGACACTTGGGAGGCGTCACCTCTGCCAGTGAATGCCACAGTTGGTGGCAG
         GTCTGTGGCAGGTGGTCGGGTCCTAAAGTCCAGATCTTGCTGTTGTTTCAAGTGATGCTC
         TGGGTGGGGGAGGAGCTGGATGGGAGAAGCCAGTGGGCGGGAAGCCTTTTTGCTGCAGGA
         [C,T]
         AGACCCTCCCACTCCAGATGACCTAGTGGCCCCTCACTGAGCCAGAAGTCCCTGTGGTGT
         GGGTGTCATGAGGTCATGTGAGGCCAACCGCCCTCCCCTGGGATGAGGCTGAGTTGGTGG
         AAGCTGATGTGGTTGTGAGGGGCTGGTGACCCTGGCTTAGGGTTTGCTGCAGGGCGGGGA

FIGURE 3Y

```
        GTCTGAGCTGGGCTGATGGTGCCATGACTGATGCGGGATGGACTACTTGCTTTCCTATGC
        TCTTGCTTAATTAGCCCTTTCCAGGCTGACTCACCCACAAGCCAGCCAAGCCAACAGCCA

39160   GATCTTGCTGTTGTTTCAAGTGATGCTCTGGGTGGGGGAGGAGCTGGATGGGAGAAGCCA
        GTGGGCGGGAAGCCTTTTTGCTGCAGGACAGACCCTCCCACTCCAGATGACCTAGTGGCC
        CCTCACTGAGCCAGAAGTCCCTGTGGTGTGGGTGTCATGAGGTCATGTGAGGCCAACCGC
        CCTCCCCTGGGATGAGGCTGAGTTGGTGGAAGCTGATGTGGTTGTGAGGGGCTGGTGACC
        CTGGCTTAGGGTTTGCTGCAGGGCGGGGAGTCTGAGCTGGGCTGATGGTGCCATGACTGA
        [T,C]
        GCGGGATGGACTACTTGCTTTCCTATGCTCTTGCTTAATTAGCCCTTTCCAGGCTGACTC
        ACCCACAAGCCAGCCAAGCCAACAGCCAGGGCTCCAGTTCAGGGACTAGCCCTCAGCTGA
        CTGGTGAAGCCTTTGTGTTTATTTCTCTGTGTTCTTTTAGGAAAACATTGCTGGGGTCTT
        CAACAAGTCCTGTGCCATCTCCTACACGAGCTACAGGAACATCTTCCCCATCTGGGCCCT
        CGGCCGCTTCTCCCAGCTGTACCCTGAGAGAGCCCTTGCTGGCCACCCCTGAGAACATGC

40405   TGTCATGGAGCCGCATACCCCTGGGTTGTGAGCTCGCCTGCATATGCAGGGTCTGTCATG
        GAACATCCCAAGTCTGTGCAGCAGGGGAGCCCCATGCCCCTGGGACATGAACCCACCTGC
        GTGGAATGCTGTTTGTGAGGTGTCTACAGGGTTTATAGTAGTCTTGTGGACACAGAAATG
        CACAGGGGACACTTACGGACACAGAAATGCACAGGGGAGGCCGAGCATAACCAGGGGTGA
        GGGGCAGGCAGCAGTTGTAGTTACTGCCGCGGGGCACTGCTATGTGCAGGGACAGCCAGC
        [G,A]
        CCCAGCCCATCACCACTCCCTGGGCTGGCTGGCAGGTATGGCACCCTGGGAGCCCGGCAT
        ATACCCAGGGCACCCCTACGGCTGCCGCCAGTCTCATGCCCAGGTGGGTGCTCTGGGCTG
        GAGCGAGGGCCAGGTTTTGGGCCGAGGCTTCCCCAGGCAATCCTGTGAGCTCCCTTCTAG
        CCTCTGACCCAGTCTGGTCTGGCTTGCATGGATGTAGGGCTTGGGGTGGGAAGTTCAGGT
        CCTGGCTTTGCCTTTGCCTGATGTGGATGAGCAGCTCACATGCTCAGGGCCACCTGAGAC

40794   CAGTCTCATGCCCAGGTGGGTGCTCTGGGCTGGAGCGAGGGCCAGGTTTTGGGCCGAGGC
        TTCCCCAGGCAATCCTGTGAGCTCCCTTCTAGCCTCTGACCCAGTCTGGTCTGGCTTGCA
        TGGATGTAGGGCTTGGGGTGGGAAGTTCAGGTCCTGGCTTTGCCTTTGCCTGATGTGGAT
        GAGCAGCTCACATGCTCAGGGCCACCTGAGACTGTCACTGCTCTCCCCTGGCTACTGGGA
        GGAGTCACTGAGAGCTTCGTTACCCCTGCTGCCTTGCCCAGGGCACACCCTATACCTCCT
        [C,T]
        ATCTGCTCTTCCCCTCCCTGCCGCCTTCTGGGCAGGTAGCAGTCCCTGGCCTCTCCCCCT
        GGCTGATCACTCTCCCTCAGGCAGTGGAGATCTGCGTCTGGACACCCTCAGATCCTGTCA
        TTGCCTGCCCAGAGTCCTTCAGGGGCACCCCTCTGCCTTGGTGTGCGGTCCAGGGCTCTC
        ACCCAGGTGCCGCACCCTCTGGGGTCTTCTGTCCAGCTCCCTTGCCCCATGTGCTGTCAC
        TGACTCTCCTTGGGACTCGCCTGCCTGCTCAGAGCCCTGCAGGGCTTGGTCAGCTGCCTG

40961   GCCTGATGTGGATGAGCAGCTCACATGCTCAGGGCCACCTGAGACTGTCACTGCTCTCCC
        CTGGCTACTGGGAGGAGTCACTGAGAGCTTCGTTACCCCTGCTGCCTTGCCCAGGGCACA
        CCCTATACCTCCTCATCTGCTCTTCCCCTCCCTGCCGCCTTCTGGGCAGGTAGCAGTCCC
        TGGCCTCTCCCCCTGGCTGATCACTCTCCCTCAGGCAGTGGAGATCTGCGTCTGGACACC
        CTCAGATCCTGTCATTGCCTGCCCAGAGTCCTTCAGGGGCACCCCTCTGCCTTGGTGTGC
        [A,G]
        GTCCAGGGCTCTCACCCAGGTGCCGCACCCTCTGGGGTCTTCTGTCCAGCTCCCTTGCCC
        CATGTGCTGTCACTGACTCTCCTTGGGACTCGCCTGCCTGCTCAGAGCCCTGCAGGGCTT
        GGTCAGCTGCCTGTTCAGTGTCAACACTTCCCTGCACATCTTAAAACTGGGCTTTATTTT
        CGCTGAAGGAACTGTGTTGGGACCCTTGACATCTGTCAGGTTTGCACATGCTGTTTTTTT
        TTCTCAGCCCACGTGTTCTCCCCCACGTGGGGTAGCAGCAGGACAGACAGTGAATCACAG

41891   AGGGAAAACAGATATTTTAAGAGATAATAGCATAGCCTATTTTAATATGTTTTAAAGGCC
        ATAAGCATATCCAGGAAGATAAATAAACGTGATACAATGTCCACATAGGAGGAACTTTCT
        TTCACTGCATTGTTTTCCTTCACAGTGGCCTTCAAGTCACAGGACGCAGCGATTCCCTGC
        CCTCTTCGGTGTTATTACACAGGCAGGACTTCAGTGTCAGTATCCCTGCCTTCAGTCTTC
        TTTAGAAATCACATCTGTGTTCAATCCATTGTTTAGAGGGAGTGTATTTTTCCTGTTCCA
        [C,T]
        GAAGAGGACTTTTTGTTCACAATTGGATCACAATGCAGAGGAGTCTGTTCCTCCCCCGTC
        GGCTTCTCGGTGCTGGGAGGGTGACCTGTCCCAGATGACTCATCACCCTGACATGCTCTT
        GACAAAGGACACCACCAAGAGGAGATGGCAGCTGTACCGGTGCAGCCTCTGTCTGAGGGG
        GATATTTGCCTCAGTGTGATTAAAAATCAGTCATGAAAGATTTTTGAATTCAGATTATTT
        TTATCAGGAACAGATTTTGAACATCCTGAAATCTTTTCCCTGGCATCATATTAGGTTTTC
```

FIGURE 3Z